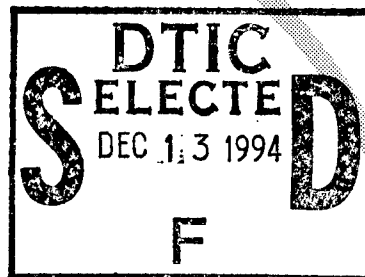DOT/FAA/CT-94/108

FAA Technical Center
Atlantic City International Airport,
N.J. 08405

Review of the Literature Related to Screening Airline Passenger Baggage

DTIC
SELECTED
DEC 13 1994
F

October 1994

Interim Report

This document is available to the public
through the National Technical Information
Service, Springfield, Virginia 22161.

19941206 055

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-94/108		

4. Title and Subtitle	5. Report Date
REVIEW OF THE LITERATURE RELATED TO SCREENING AIRLINE PASSENGER BAGGAGE	October 1994
	6. Performing Organization Code
	ACA-400

7. Author(s)	8. Performing Organization Report No.
Ronald J. Lofaro, Ph.D., Gerald D. Gibb, Ph.D., Daniel J. Garland, Ph.D., Mr. Patrick C. Guide, Mr. James S. Baker, Mr. Mark St. Laurent	

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405	
	11. Contract or Grant No.
	FAA Grant 93-G-037

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
U.S. Department of Transportation Federal Aviation Administration Assistant Administrator for Civil Aviation Security Washington, DC 20590	Final
	14. Sponsoring Agency Code
	ACS-1

15. Supplementary Notes

16. Abstract

The purpose of this research is to identify the underlying abilities and traits required for successful performance as an airline passenger security screener. The primary focus of this effort is to develop a validated selection battery and protocols that can be easily administered across a wide dispersion of organizations in the aviation security industry. Aside from the obvious costs associated with recruiting, selecting, and training replacement employees, there is likely to be a detrimental impact on the effectiveness of airline passenger screening when a substantial percentage of the workforce are novice workers. This research program seeks to identify the causes of employee satisfaction/dissatisfaction and develop appropriate interventions to curb existing retention problems. A final objective of this program is to establish criteria that discriminates levels of effectiveness among screeners. An effective and reliable performance measurement system is required for this effort, and can provide useful criteria for other programs responsible for the establishment and validation of screener training, evaluation of advanced hardware, and implementation of screener performance evaluation systems.

DTIC QUALITY INSPECTED 8

17. Key Words	18. Distribution Statement
Airline baggage screeners, personnel, performance, selection, airline passenger security screener/pre employment inventory, APSS/PI, training, turnover, vigilance, performance measures, human factors	This document is available to the public through the National Technical Information Service (NTIS), Springfield, Virginia, 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	46	

Form DOT F1700.7 (8-72)   Reproduction of completed page authorized

TABLE OF CONTENTS

Accesion For		
NTIS CRA&I	☑	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

LIST OF ILLUSTRATIONS

LIST OF TABLES

EXECUTIVE SUMMARY

The purpose of this research is to identify the underlying abilities and traits required for successful performance as an airline passenger security screener. The primary focus of this effort is to develop a validated selection battery and protocols that can be easily administered across a wide dispersion of organizations in the aviation security industry. We expect a wide range of products to emerge from these efforts that can include performance tests, screener interviewer training, and interviewing protocols.

First year efforts however also indicated that a severe employee turnover problem exists. As a result, initiatives to identify factors contributing to job satisfaction and career retention were incorporated. Understanding the causes of employee turnover and developing viable solutions became an integral component of this project.

A series of three reports contain our first years' efforts and focuses on the literature review, acquisition and development of subject matter expert (SME) data, and job task analyses (JTA). The first report presents the background and literature review. It contains a summary of prior work related to the selection of screeners, selection work in related occupations, application of concepts in vigilance research, human factors problems, and issues in checkpoint design. The second report describes the results of the job task analysis and concepts related to the development of a prototype selection test battery. The final report in the series examines the issues related to employee turnover. This report discusses the Delphi analysis used to identify job satisfaction and dissatisfaction factors.

SME data were generated through interviews during the JTAs and modified Delphi techniques using 115 screeners. Concurrent with the development of the SME database, JTAs were conducted at various security sites across the nation. Considerable attention was given toward identifying a representative sample of each site. Integration of these results will lead to the development of a taxonomy of abilities and traits that have potential as predictors of screener performance. In addition, causes of self-initiated employment termination will identify issues and factors that have to be addressed to retain effective airline screening personnel.

Second year efforts will primarily focus on the initial validation of a prototype selection test battery and the development of intervention techniques to improve career longevity. It is expected that a considerable portion of this phase will be devoted to validating constituent tests with strong psychometric characteristics and establishing reliable performance criteria using false image projection technology. Data relating to critical psychometric properties (i.e., differential stability, task reliability) will be collected within a university laboratory environment. Concurrently, Klein and Associates, in conjunction with our efforts, will develop and validate a performance measurement system that discriminates effective screeners from those who demonstrate performance problems. False image projection during on-line operations has been chosen as the most pragmatic and useful performance criteria to assess the predictor variables in the validation efforts. This methodology offers the best approach to distinguishing levels of performance among screeners. Field testing efforts during the second year will concentrate on

initial validity studies, using new hires shortly after the company application process has been completed. New hires will be administered the prototype tests during the application and pre-employment process, but before training is initiated.

Several intervention programs to mitigate turnover rates will be developed during this phase of the program. As intervention programs are implemented at various facilities, average employee tenure rates, employee turnover, passenger/aircrew complaints and other criteria will be tracked to assess the impact of the interventions. In conjunction with this effort, we will be conducting a nationwide survey to scale the items that resulted from the Delphi analyses. Scaling techniques were selected to develop relative weights for each of the satisfiers and dissatisfiers that emerged during the Delphi workshops. In addition, the data will assist in assessing the generalizability of the findings from the Delphi analyses.

Additional data will be collected at select sites using the Likert Characteristics of Organizational Climate survey to assess the divergence between how screeners perceive their current work environment and how they would like it to be. These data will also be used in developing intervention programs designed to reduce the high turnover rates.

1. BACKGROUND AND SIGNIFICANCE.

The safeguarding of civil aviation against hijackings and/or sabotage has been a grave concern to governments around the world. Such malicious acts have had profound detrimental effects on civil aviation (e.g., the bombing of Pan Am flight 103 over Lockerbie, Scotland). The cost of such acts, when considering the loss of lives and disruption to air carrier operations, is unacceptable.

Since the inception of the passenger and carry-on baggage screening program in 1973, over 11 billion passengers and their carry-on items have passed through airport security checkpoints. According to the Federal Aviation Administration (FAA, 1991), this has resulted in the detection of more than 45,600 firearms and over 20,150 arrests.

Periodic inspections evaluating the effectiveness of airline passenger security screening are reported to average about 90 percent for success in detecting "test weapons" submitted by FAA inspectors posing as passengers. However, a 1987 General Accounting Office report (GAO, 1987a) found that the nationwide success rate in such tests was only 80 percent. More importantly, their analysis found tremendous variability between security facilities in detecting "test weapons." One site had a "hit" rate (i.e., success rate) of only 34 percent. The GAO report clearly demonstrates the lack of standardization and effectiveness among many security facilities.

In a report by the President's Commission on Aviation Security and Terrorism (President's Commission, 1990), the commission members noted the importance of adequate FAA guidance in the selection of security screeners. The commission was critical toward the FAA with regard to how little attention was paid on recruiting and motivating security personnel.

To date there has been little emphasis placed on the selection of security screeners. The Air Transport Association (ATA), in recognizing this shortfall, provided input and limited funding for the development of a 32-item questionnaire to assess specific attributes (e.g., attentiveness, dependability, attitude). The breadth of this instrument is quite limited however, and the validity is unknown. An analysis of this instrument is provided in a subsequent section of this report.

These concerns lay the foundation for this research. Standardized screening and selection of security personnel provide a means for hiring a high quality workforce while reducing attrition of those suited for the profession. In addition, selection of potential personnel can be accomplished using research-based algorithms wherein the probability of a successful hire can be estimated. A uniform approach to selecting security personnel, based on aptitude and ability, further offers the potential to developing standardized training programs. This would be possible since training objectives and content could be directed at a more homogenous candidate population with the requisite baseline abilities and skills.

1.1 PURPOSE OF THIS RESEARCH.

In recognizing the critical need for an effective and stable airline passenger security screening workforce, this research is directed toward developing a screening program that identifies and

staffs the nation's airline security system with the most capable and motivated individuals. The success and deterrent potential of that system is primarily dependent on the personnel who operate it. An effective selection process for security personnel that is standardized nationwide will serve to enhance the overall capability and deterrent potential of the entire airline security system.

A priority that emerged during the first year's work was the critical need to address the problem of extraordinarily high employee turnover rates and training attrition rates. Aside from the obvious costs associated with recruiting, selecting, and training replacement employees, there is likely to be a detrimental impact on the effectiveness of airline passenger screening when a substantial percentage of the workforce are novice workers. This research program seeks to identify the causes of employee satisfaction/dissatisfaction and develop appropriate interventions to curb existing retention problems.

A final objective of this program is to establish performance criteria that discriminates levels of effectiveness among screeners. An effective and reliable performance measurement system is required for this effort, and can provide useful criteria for other programs responsible for the establishment and validation of screener training, evaluation of advanced hardware, and implementation of screener performance evaluation systems.

2. INTRODUCTION.

A successful national airline passenger security screening program is predicated on effective FAA guidance in the selection of screeners and in their training, as well as on the importance which individual airlines place on security. The purpose of passenger and carry-on item screening is to ensure that crimes against civil aviation are minimized and ultimately eliminated. In order to achieve this, the security team is provided with relatively sophisticated equipment that requires a thorough understanding of its capabilities; but more importantly, they must be able to operate the equipment effectively under varying workload conditions (e.g., peak periods or airport "rush hours"). As with most safety-critical systems, there is no room for system-induced or operator-induced errors.

Criminal acts against civil aviation have led to a concerted effort to strengthen the international aviation security system. In response to a series of crimes against civil aviation, the U.S. Congress mandated, through amendments to the Federal Aviation Act of 1958, that the FAA assume primary responsibility for civil aviation security. In accordance with the act, the FAA, in cooperation with airports and air carriers, is responsible for ensuring that security programs and procedures are instituted that will safeguard passengers, crew, aircraft, and airports. As part of its mission, the FAA Office of Civil Aviation Security establishes security requirements, inspects airline and airport security operations, and issues civil penalties for noncompliance with the requirements.

The Air Transportation Act of 1974 requires that all passengers and property intended to be carried in an aircraft employed in interstate transportation be screened by weapon-detection procedures or agents of the air carrier. Section 315 of the Federal Aviation Act directs the FAA

2

to prescribe regulations requiring the screening of all passengers and carry-on items for the presence of unauthorized items. Section 316 of the Act also requires regulations to protect persons and property aboard aircraft from acts of criminal violence and piracy. Federal Aviation Regulations (FARs) Parts 107 and 108 govern domestic airport and air carrier security, respectively. These regulations mandate the adoption and effective implementation of minimum security programs by airports and air carriers. These programs must include specific measures for passenger screening, protection of aircraft, and airport access controls. In general, FAR Part 107 requires that the airport operator: (a) create a security program for the airport, (b) provide controls to prevent or deter unauthorized persons from accessing the air operations area, and (c) provide law enforcement support. For air carriers, FAR Part 108 generally requires that the carrier: (a) adopt and carry out a security program, (b) screen passengers and property, (c) provide and use ground and in-flight security coordinators, and (d) prohibit unauthorized access to the airplane.

Indirect cargo air carriers (ICAC) are required under FAR Part 109 to adopt and carry out a security program approved by the FAA. This security program is intended to prevent and deter the unauthorized introduction of explosives and incendiaries into packaged cargo and mail tendered in air commerce. The U.S. Postal Services (USPS) and its military counterpart, the Military Postal Service Agency (MPSA), maintain significant jurisdictional and practical control over domestic and international airmail carried by U.S. air carriers. The FAA is currently working with the USPS and MPSA to review airmail security and screening procedures.

Air carriers are responsible for screening passengers with metal detectors and X-raying their carry-on items for weapons and explosives. Air carriers have generally elected to contract with private security firms to perform security functions. At the time of this writing, there were over 15,000 screeners employed by some 46 security firms nationwide.

The air carriers will frequently hire the security firm with the lowest bid (St. John, 1991). Nonetheless, the air carrier is held responsible by the FAA for the effectiveness of the screening procedures, and both the air carriers and the FAA monitor a security firm's performance. Air carriers are responsible for developing security plans which assure prevention or deterrence of aircraft hijacking, sabotage, and related criminal acts. The airport operators are to maintain security programs which prevent and deter unauthorized entry into air operations areas and provide law enforcement support for the screening system and overall airport security requirements (FAA, 1991).

The FAA's role in aviation security expanded significantly in 1985 with passage of Public Law 99-83, the International Security and Development Cooperation Act, with further increases in FAA involvement in 1990 with the passage of PL 101-604 (the Aviation Security Improvement Act), as a result of the downing of Pan American flight 103 over Lockerbie, Scotland. The Act required the FAA to assess the adequacy of security at foreign airports served by U.S. carriers, and the security procedures of foreign air carriers flying to the U.S.

3.   REVIEW OF THE RELATED LITERATURE.

3.1   PASSENGER AND CARRY-ON ITEM SCREENING.

The most visible aspect of domestic airline security is the screening of passengers and carry-on items. For all practical purposes, the focus of the security procedures for domestic flights is to deter hijackings and has been so since their inception in 1973. The GAO (1987a) report found that for a given six month period (July 1, 1989 to December 31, 1989), over 535 million persons were processed through screening checkpoints at U.S. airports. This process resulted in the detection of 1,464 firearms, of which 1,406 (96 percent) were detected by X-ray inspection, 30 (2 percent) by physical search, and 28 (2 percent) by use of metal detectors. In addition, 10 explosive/incendiary devices were discovered during the period including five grenades, three fireworks, one flare gun, and one tear gas device. During this six month period, 764 persons were arrested at passenger screening points for unauthorized carriage of firearms or explosives/incendiary devices.

The FAA's testing of the effectiveness of the screening process is straightforward but somewhat simplistic. The agency has historically used test items, such as simulated dynamite tied together with a large clock and attached wires, to test the X-ray equipment and the ability of the operator to detect a potentially lethal weapon. The test item is generally placed in a brief case or bag with little effort to conceal or disguise it. The briefcase or bag is then taken by an FAA security inspector posing as a passenger, to the passenger screening point and submitted to the screening process.

One would expect near 100 percent success rate in detecting "test weapons" since the FAA test objects resemble only very obvious threats and since the FAA inspectors are not allowed to hide the test object(s) into the carry-on (GAO, 1987b). Furthermore, screeners are often aware that they are being tested because FAA inspectors are well known to screeners at most airports. One investigator (Vosburgh, 1993) contends that this approach is not very practical. He noted that most screeners "... would probably not recognize a real bomb [or threat] since they train primarily to identify test objects similar to those used during airline and FAA tests" (p. 61).

The FAA reports that, nationwide, the screening process is identifying FAA test weapons slightly over 90 percent of the time (FAA, 1991). The screening system's performance has improved since 1987, when the GAO noted that the tests found an average detection rate of about 80 percent (GAO, 1987a). These percentages are based on 2,419 screening tests conducted at 28 major airports (i.e., those that screen over 2 million persons annually) from September, 1986 to December, 1986 (GAO, 1987a).

Airline passenger security screeners rely on relatively sophisticated equipment consisting of walk-through metal detectors and X-ray inspection systems to screen passengers and their carry-on items. Hand-held metal detection devices are also used, but primarily as backup support for walk-through detectors. In addition, screeners may require physical searches for items in suspicious carry-on baggage. Each component of the process—X-ray, metal detector, and physical search—is periodically tested by the airline and FAA. While there have been some

technological improvements to screening equipment, the process generally operates the same today as it did when implemented in 1973. Nonetheless, some industry critics (St. John, 1991) report that the airlines put their trust in "... metal detectors and X-ray machines—spending millions on equipment and little on the people who operate it" (p. 85). Regardless of the amount spent and the justification for such advanced equipment, Hughes (1993) noted that a key conclusion of the latest National Research Council study is that there is no single detection technology available today that can provide a high probability of bomb or weapon detection.

There is an effort however, in the FAA's Research and Development Program, to enhance and automate X-ray systems used in the screening of passengers and carry-on items. Several manufacturers are independently showing significant innovation in extending X-ray technology to identify specific threats. Work is underway to improve concourse X-ray system performance by concentrating on the development of automatic pattern recognition software and hardware. Integrated into current X-ray detectors, automatic pattern recognition systems could alert the operator to suspicious items in luggage by graphically highlighting the suspect item.

3.2   AIRLINE PASSENGER SCREENING PERSONNEL PERFORMANCE ISSUES.

On December 21, 1988, an explosive device was detonated aboard Pan American (Pan Am) Flight 103 en route from London to New York resulting in the deaths of all persons on board and 11 persons on the ground in Lockerbie, Scotland. As a result of this tragedy, the President's Commission on Aviation Security and Terrorism was formed on August 4, 1989, by Executive Order 12686 (President's Commission, 1990). The Commission's objective was to conduct a comprehensive study and appraisal of the practices and policy options to prevent terrorist acts against civil aviation.

In a recent study, Vosburgh (1993) noted that a combination of errors on the part of the carrier and their security contractor contributed to the destruction of Pan Am Flight 103. Vosburgh cited some of the contributing factors; they included: "... passenger screening; baggage handling, especially matching passengers and their bags; [and] almost complete reliance on X-ray equipment to find explosive devices ..." (p. 5).

The importance of having a consistent set of selection and training standards for airport security was demonstrated by the Pan Am 103 investigation. The investigation suggested that the security deficiencies found could be connected to breakdowns in airline security performance. In addition, the investigation found that Pan Am security personnel failed to properly screen 38 passengers at Heathrow airport (President's Commission, 1990).

During the late 1960s and early 1970s, U.S. air carriers had experienced a sudden and unexpected increase in domestic aircraft hijackings (see figure 1). The FAA responded with a coordinated and extensive civil aviation security system whose primary deterrent was passenger screening using metal detectors and X-ray machines. Figure 1 illustrates the effectiveness of these security procedures by comparing the hijacking statistics before and after the onset of passenger screening implemented in January of 1973.

Another study briefly discussed in "Study Tallies" (1991) examined hostile acts against civil aviation between January 1980 and December 1990. During that 11-year period, the investigators uncovered 304 hijacking incidents—one every 13 days. Foreign air carriers however, encountered the majority of these hijackings.
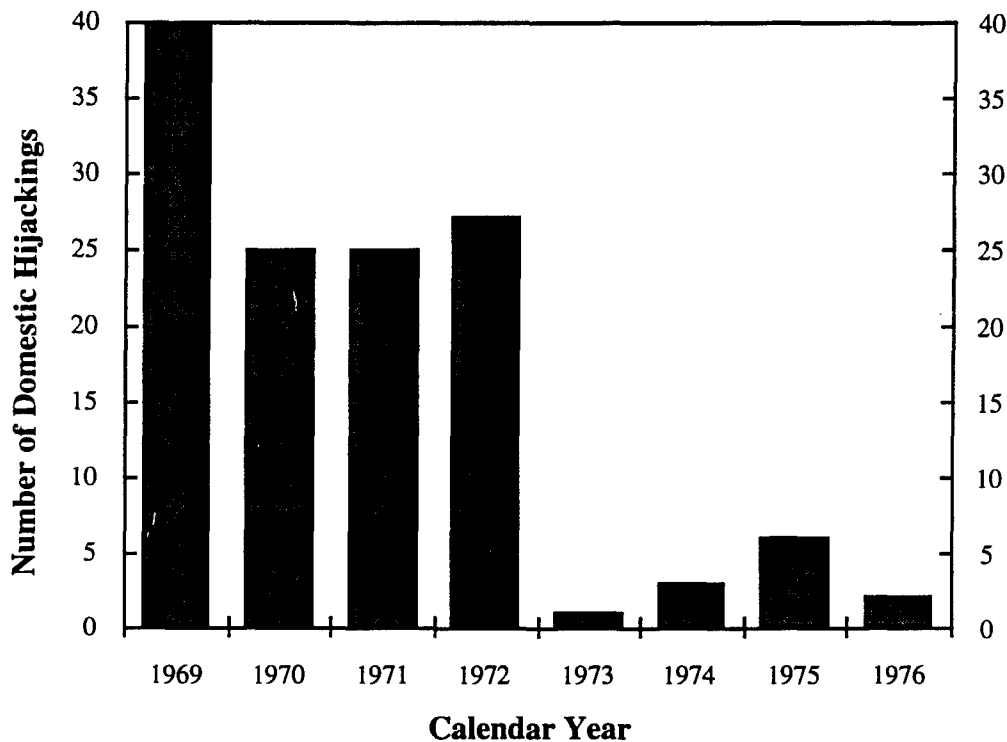
FIGURE 1. THE EFFECTIVENESS OF PASSENGER SCREENING. (ADAPTED FROM "SCREENER TRAINING PROGRAM INSTRUCTOR'S GUIDE" BY AIR TRANSPORT ASSOCIATION OF AMERICA (ATA) IN COLLABORATION WITH THE FAA.)

Aeroflot had suffered the most with 24, followed by LOT Airlines with 16, and Iran Air with 15. Although hijacking incidents do not appear to be an immediate threat for U.S. operators, St. John (1991) stated, "As international security tightens, the United States, with its lack of safeguards at airports, becomes more inviting to terrorists" (p. 85).

Oster, Strong, and Zorn (1992) reported that the risk of a terrorist incident or sabotage act is always present and "... is not confined to the middle east and other trouble spots in the world" (p. 143). Supporting this position, Oster et al. investigated the number of explosions aboard aircraft by geographic region between 1950 and 1989. They found a total of 72 incidents, with Asia experiencing 20 (28 percent), North America with 18 (25 percent), Western Europe with 15 (21 percent), the Middle East with 9 (13 percent), and the remaining 10 in South America, Africa, and Eastern Europe.

The FAA has initiated several programs to counter terrorist threats. The FAA sponsors aviation security training programs at the Transportation Safety Institute (TSI) in Oklahoma City, Oklahoma under the authority of Section 316(c) of the Federal Aviation Act of 1958. The current programs relating to aviation safety and security are administered by the Aviation Security Division (ASD) of TSI. The FAA identifies in its security program the core requirements and guidance for the initial, recurrent, and on-the-job training of airline passenger security screeners at domestic airports. However, several task forces and studies have found that the quality of this training varies widely among the airlines.

The ATA has recognized the need for improvements in the selection and training of screening personnel. The ATA encourages air carriers to conduct tests for screeners on a regular basis. However, these tests use the identical testing objects used by the FAA inspectors. In 1989, there were 56,000 tests performed by the air carriers with a reported 96 percent detection rate (ATA, 1990). However, the validity of these tests remains questionable. Often screeners recognize the air carrier personnel who conduct the tests and are alert to the likelihood that a test will be conducted. Actual threats typically occur without warning and without salient cues present. In order to further improve on this performance, the ATA (1990) has developed a pre-employment inventory for selecting security screeners. However, this test instrument—the Airline Passenger Security Screener Pre-Employment Inventory (APSS/PI)—has received limited implementation and its validity is also unknown.

The ATA has developed programs for hiring and retaining quality personnel at security positions. These include background checks, aptitude tests, psychological screening, English language proficiency, and wage surveys. The FAA has since adopted many of these guidelines. In addition, personality profiles have been developed by the ATA for the security field, reporting that nearly two-thirds of security personnel applicants can be disqualified because they do not fit these presumed profiles (ATA, 1990).

In 1978, following unsatisfactory test results of the security screening process, a task force of the FAA and airline security personnel studied ways to improve performance at passenger screening checkpoints. This task force's report, referred to as the "Human Factors Study," recommended several actions which were eventually endorsed by both the FAA and the airlines. For the most part, these recommendations focused on the personnel-related aspects of the process such as high employee turnover rates, low pay, and inadequate training. Although the FAA and the airline industry endorsed the study's recommendations, the air carriers have been slow in fully implementing them.

In a 1987 GAO report to the Secretary of Transportation (GAO, 1987b), it was reported that an investigation of screening processes at six major airports found that many of the problems that were addressed in the 1978 human factors study, in fact, still existed. For example, security firm managers reported that screeners were still being paid at or near minimum wage and that low pay contributes to high-turnover rates—in some cases, about 100 percent annually—and further resulted in problems in hiring capable people.

The responsibility for establishing an effective and integrated aviation security program is shared by the air carriers, airports, Federal, State, and local governments (FAA, 1991). In an unpublished study, Thaher (1991) reported that the FAA spent $10 million in 1990 on the development of new passenger screening, explosive detection, and baggage inspection technology. However the aviation security industry, more specifically the airlines, have been accused of a penny-pinching attitude toward security (St. John, 1991). Furthermore, it has been suggested that only 65 cents per passenger ticket goes toward security, in contrast to the passenger facility charge (PFC) which can top out at $3.00 per passenger ticket. These limited funds obviously affect the quality of the equipment obtainable as well as the quality of personnel selection and training. "The last thing airlines want to consider is spending more money for security. What they do not seem to realize is that security must be a component part of their service, just like on-time flights, safety, and passenger satisfaction" (Vosburgh, 1993; p. 6).

The International Civil Aviation Organization (ICAO) has also identified a number of areas in aviation security requiring attention (Coninsx, 1993). These include: low wages, lack of career prospects, lack of challenge in the job, lack of authority, poor working hours, pressure from the airlines to perform during peak hours, fear of possible health effects (e.g., bomb detonation, radiation exposure), penalties for poor performance, and lack of sufficient knowledge to handle dangerous situations (e.g., armed passengers, explosive devices). We must first create better working conditions and better salaries in order to attract suitable personnel. When appropriate, we must also reward employees for their efforts, loyalty, and achievements.

It is impossible to overstate the importance of selecting high-caliber security personnel. Wage constraints cannot be accepted as a valid reason for hiring a substandard security force. Adequate employee screening is a vital prerequisite to an effective long-term security screening operation. After employees are hired, their performance should be realistically critiqued with regular performance evaluations as done with other aviation occupations (e.g., pilots, air traffic controllers).

3.3 SELECTION AND HIRING PROCEDURES.

The selection process of those persons who have the necessary attributes for target detection is relatively unexplored. Selection procedures typically assess candidates in terms of test scores that serve as predictors of job success (Atkinson, 1973). The purpose of personnel selection is to identify the most suitable applicants, train them adequately, and ensure they retain their employment within the organization. Who are the individuals that are best able to detect potential threat items? Who are the individuals that are better able to maintain attention over a long duration of time? Are there similar personality characteristics of the good performers? If so, what are they? Can airline passenger security screeners be identified as successful performers through a series of tests, attributes, and/or biographical information?

Coninsx (1993) noted that a precise personnel specification, or candidate profile, is required for an efficient recruitment process; however none exist for airline passenger security screeners. In order to be able to carry out the job to the required performance standards, the personnel specification identifies the most important areas of knowledge, skills, and personality

characteristics required. Drawing up these specifications is not easy. They often include requirements on physical abilities, general intelligence, interests, disposition, and motivation (Coninsx, 1993). In addition, the requirements need to be adapted to the evolution of the technology utilized. For example, recent advancements in X-ray displays utilize color to detect various hazards, thus those screeners who can not discriminate specific colors should be screened out of the pre-employment selection process. This, however, raises some legitimate concerns for those screeners who have been on the job for years prior to implementing these new displays.

Some investigators (Pierce, Crumley, and Clifford, 1991) report that few guidelines exist for selecting personnel who are best suited to perform vigilance tasks (e.g., X-ray screening). Pierce et al. suggest that there is potential for improving target detection performance by creating a scale which can accurately identify those candidates who are vigilance maintainers. In addition, other investigators (Boff and Lincoln, 1988) emphasize the need to develop tests that could correlate vigilance performance with membership in a class or group.

Even the most advanced security equipment employing the latest technology can become useless in the case of human error, human negligence, and personnel inefficiency. Some investigators (Coninsx, 1993; Drury and Fox, 1975; Funke, 1979; "Keeping the operator," 1992; Vosburgh, 1993) purport that the human has proven to be a weak link in inspection systems. Nonetheless, detecting targets in low signal-to-noise environments where search time is limited can, at times, be a difficult task, particularly when the signals (i.e., targets) are intermittent, unpredictable, and infrequent.

Screeners must always maintain a high level of alertness so that they do not miss any potential threats. Sustaining high levels of attention is critical for such related occupations as air traffic control, nuclear power plant monitoring, industrial inspection, and piloting an aircraft (Edwards, 1990). For example, detecting the components of a bomb (i.e., timing device, initiator, power source, explosive) as seen through an X-ray requires attentive behavior. The screener must then, during the brief presentation of the image, identify the components and determine if they are part of an obvious threat. The components must be identified for what they are, and then associated with the explosive device. Since the parts, whether explosive or weapon elements, can be spread out over several bags, the process can be quite difficult (Vosburgh, 1993).

3.3.1 Current Practices in Passenger Screener Personnel Selection.

The position of airline passenger security screener is relatively new, having been established as recently as 1973. Unlike other aviation personnel (e.g., air traffic controllers, flight deck crew, and maintenance technicians), comparatively little attention has been devoted to developing standardized selection protocols for this occupation. This can be attributed to a number of factors including: (a) relatively low salary scales; (b) lack of lengthy and costly training required; (c) minimal education and technical requirements for entry to the field; (d) no similar military counterpart; (e) lack of job complexity; (f) traditionally high turn-over rates; and (g) the decentralization of hiring resulting from a host of private concerns employing these individuals. Private organizations do not have the resources, or the expertise to develop specialized selection

9

programs. Consequently, there is a marked paucity of personnel selection research that exists for this particular occupation.

Indeed the only effort aimed at the development of an airline passenger security screener selection system in the United States is a series of studies commissioned by the Air Transport Association (ATA). In 1988, the ATA contracted with an independent consultant group (hereafter referred to as the contractor) to develop selection protocols that could be widely used by the ATA member organizations.

3.3.2 Previous Research on the Selection and Hiring of Airline Passenger Security Screeners.

3.3.2.1 Critique of the APSS/PI.

The contractor initially used a concurrent validation technique to construct their selection instrument, referred to as the Airline Passenger Security Screener/Pre-Employment Inventory (APSS/PI). The instrument focuses primarily on personality traits believed to be related to job performance. These traits were derived from a review of available job description information and interviews with security screening experts. The basis of the contractor's work relied on administering the Criterion Inventory Series (CIS) to over 200 screeners working in five geographically dispersed airports across the country. Those particular airports will not be discussed in this report. Performance on the predictor instrument was compared with several operational job performance measures. These measures included: (a) FAA testing errors, (b) airline testing errors, (c) security firm testing errors, (d) supervisor ratings, and (e) unexcused absences or incidences of reporting late to the job site. The measures were chosen by ATA Security Committee members and managerial personnel at the security firms for their availability (i.e., screener personnel files), their operational significance, and ease of measurement.

The contractor found a relationship between several of the predictor items and both objective job performance measures (testing errors) and personal reliability measures (absenteeism/lateness). No relationship was found between supervisor evaluations and the predictors or the objective job performance measures. The analysis of these data was used to construct the APSS/PI. The APSS/PI is then a modification of the CIS and includes within the items an "honesty" (reliability) scale. Elevated scores on the APSS/PI indicated a "risky" applicant and generated a "not recommended for hire" classification. Those who scored high on the APSS/PI were found to make significantly more screening errors, have higher absentee rates, and have poorer job performance ratings.

In 1990, the contractor conducted an evaluation study with an additional group of 211 screeners. Although several new test sites and participants were added to the study, it essentially remained a concurrent validation similar to the 1988 effort. APSS/PI scores were compared with a number of testing errors and job tardiness measures found in the personnel records of screener participants for the preceding six month period. Months of employment and supervisory ratings (disciplinary problems, communication skills, attentiveness, dependability, and attitude) on a 5-point scale were also recorded. As previously found, a significant relationship was reported to

exist between APSS/PI scores and objective performance measures. Again, there was no relationship between APSS/PI and subjective evaluations provided by supervisory personnel.

There are a number of serious problems with this developmental effort. Of considerable concern is the use of concurrent validation to establish the predictive utility of the APSS/PI instrument. In both the 1988 and 1990 evaluations, APSS/PI "scores" were compared against historical data in the employee's personnel records. The use of this approach is highly questionable, particularly in the 1990 evaluation. Given the nature of the criteria and the use of periodic screening performance checks, why was there not a predictive validation conducted in 1990? New screeners could have been administered the APSS/PI, predictions generated, and their performance evaluated through the established techniques. Concurrent validation techniques do not assure predictive validation. The use of incumbent screeners, as opposed to applicants, further introduced the possibility of restriction-in-range in the sample. Data from the CIS for the 1988 sample indicated only relatively moderate variability among individuals, for example.

The investigators further falsely assumed that the incumbent screeners had all the requisite traits and abilities, to the exclusion of the possibility that others with different profiles might also be successful. The use of concurrent validation techniques also fails to account for most new hires leaving the career field. This is an important concern given the typically high turn-over rates.
The reliance on an existing, historical criterion with questionable psychometric characteristics is also problematic. A more desirable approach should have included a standardized criterion that could have been developed and administered by displaying "targets" embedded in a video tape constructed for this purpose. This approach would have eliminated any possible contamination from expectancy bias during the evaluations.

In these efforts concurrent validation was not a justifiable or suitable substitute for initial and cross-validation of the instrument. The studies are exploratory at best and do not constitute an acceptable validation as dictated in the 1978 employee selection guideline as the contractor maintains. A major issue regarding the development and validation of the APSS/PI centers on the sample selection procedures. There is some indication that part of the sample used to select the aptitude items for the APSS/PI from the CIS was also part of the sample used to validate the final scale. The investigators reported that some of the same sites and security firms were used in both studies, but no information is provided regarding if, and how many, subjects were part of both analyses. Constructing and validating a selection instrument with the same individuals is not an appropriate method.

The development and content of the APSS/PI raise numerous issues as well. The instrument is apparently focused on personality traits alone and excludes the use of other possible domains of predictors (e.g., motivation, perceptual skills, cognitive abilities) that may contribute to screener performance. The limited scope of human characteristics that the APSS/PI measures would severely curtail its ability to successfully predict performance. The most obvious omission was a failure to even consider the use of a work sample test. The contractor appears to have operated under the assumption that personality measures alone would have accounted for 100 percent of the variance in performance.

11

Exceedingly weak evidence is offered for using strong analytical methods to establish the predictor variables or their construct validity. The study relied almost exclusively on a limited sample of subject matter expert interviews and U.S. Department of Labor descriptions to derive possible predictor variables. Some indication was given that the investigators simply identified personality characteristics that would lead to negative performance (i.e., errors, job tardiness) with no mention of any type of analysis. Additionally, no information or description of how these data were collected, interpreted, and analyzed is presented. The basis for test development or selection of possible predictor variables was inadequate for this effort. Strong selection programs typically begin with a complete job or task analysis to define the underlying traits and abilities. The derivation of a taxonomy of human characteristics to perform as an airline screener is almost devoid in these studies. This consequently led to the development of a narrowly defined selection instrument.

The origination of the APSS/PI is itself problematic. While the validity and reliability of the CIS is stated to exist, no data are provided to support this. The instrument is virtually unknown and almost no description of its structure or development is offered. One is left wondering exactly what personality dimensions are assessed with the instrument.

The refinement of this instrument to the APSS/PI becomes even more puzzling. How were items selected from the CIS to establish the APSS/PI and why? Furthermore, there are no data to: (a) establish the reliability of APSS/PI; (b) demonstrate the construct validity of the scale; or (c) establish the accuracy, reliability, or validity of the embedded lie scales. No information was ever presented on how the "reliability" (lie) scale items were selected, or how these were even used in the validation study. Were subjects eliminated from the validation study if they exhibited elevated reliability (lie) scale scores, and what were the cut-off values? The performance data presented reveals various degrees of freedom among the different variables indicating some subjects were deleted from analysis. It is not known why were these subjects deleted.

The APSS/PI has been purported to measure the personality traits requisite of a successful screener. However, only a global score and a "lie" score are generated by the APSS/PI instrument. No data are presented that indicate the multi-dimensional nature of the several personality traits that were initially put forth as important. Nor was any consideration apparently given to the possibility that individuals would exhibit differing degrees of personal traits, as the scale provides no value to each purported personality dimension. In short, no measurement or evaluation of individual traits was given. It is questionable if the APSS/PI is actually a composite measure of personality at all given the marked paucity of data available. The personality dimensions that are purported to be embedded in the scale are not identified, described, or delineated by the contractor investigators. Considering the limited nature of the APSS/PI (32-items), it is unlikely that sufficient items even exist as markers to adequately assess several personality dimensions, particularly since only dichotomous responses (agree versus disagree) are possible.

Finally, given the shortfalls of the CIS and APSS/PI, why weren't established, well documented, and researched scales used to assess the personality traits of interest? The investigators provide no rationale for beginning this line of research with the CIS except for gratuitous arguments

against the lack of "customized honesty scales" that were not designed for the airport screener population. Yet no data or argument is provided to demonstrate that airport screeners are different from the general population.

An additional area of concern is the use of factor analysis and the development of composite criterion measures. Factor analysis is a data reduction procedure used to define the underlying dimensions of a set of variables. It is not clear why the investigators chose to enter three sets of testing errors, an absenteeism variable, and the five subjective supervisory variables into the analysis. As expected, the testing error measures clustered together into one "factor" while the supervisory ratings defined a separate distinct factor. Testing errors was a simple dichotomous measure of target detection versus non-detection. In this environment where the detection of "targets" is the key job performance measure embedded within identical task parameters, only the organization performing the administration of the performance check varied (i.e., the FAA, security company, airline).

There is no justification to assume the three measures were not related. Similarly, it could be expected that the five supervisory ratings would also define a factor given the nature of the scale used and the wealth of information in the literature regarding supervisor ratings and the halo effect. There was no apparent need to identify an underlying factor, nor was one fully described.

The troublesome aspect of the factor analysis was the use of these results to construct composite measures. One composite measure of objective job performance summed the total of all FAA, security company, and airline testing errors into a total testing errors measure. However, the data to form this composite measure were historical records over a 6-month period using a simple scale that recorded the number of actual screening errors. One value on that scale includes a category for six or more errors. Simple summation of error frequency counts across these scales would cause a restriction of range problem for the criterion by effectively building a ceiling effect into one end of the error criterion. This is caused by equally weighting points on a scale which were not in fact equivalent. Of greater concern is the failure to use a mean measure of errors. The base number of job performance checks is not provided leaving one to question if there were a different number of performance checks between screeners. Whether all the performance checks themselves were comparable remains uncertain. Variability in the length of the checks, number of targets ratio, and target ambiguity would introduce considerable error variance.

The second composite objective performance measure, overall negative performance, adds the number of times late to work to the previous composite measure. The investigators offer no justification/rationale for combining these two separate sets of variables. Furthermore, it appears inappropriate to equate a screening error with arriving at work late. A simple summation procedure weighs the two variables equally, however. It is plausible that a screener who is quite proficient at target detection is consistently late for work. Yet it would be impossible to distinguish this individual from a screener who always appears on time for work but is less adept at target detection, the actual operational measure of performance. Target detection performance and absenteeism should have been treated as two separate measures. No analyses were even

13

provided to indicate the two sets of measures are related. For system performance (safety), only one of these classes of measures is actually critical.

A close analysis of the data yields some interesting findings. The coefficient of correlation between APSS/PI and the total testing errors composite score is only $r = +.163$. This accounts for only 2.6 percent of the variance. An examination of individual data indicates that within the group classified as "not recommended for hire", based on the APSS/PI (n = 18), 74 percent of the testing errors within that group are made by only 33 percent (n = 6) of that subgroup. Most errors were apparently made by only a few individuals. Four of these "not recommended" screeners accounted for 18 of the 27 errors (67 percent). The performance of these four individuals (the only screeners with more than two errors) is clearly an anomaly. In fact, 12 individuals in this subgroup made one or no testing errors.

Eighteen percent (18 percent) of the "recommended" group were also reported to have at least two screening errors. Given the comparatively small sample sizes and substantial individual variances it is not likely these data can be replicated in cross-validation. More importantly, the APSS/PI as used to separate screeners into "recommended" versus "not recommended" subgroups was ineffective in classifying screeners by error rates. The contractor offers no rationale for this inconsistency. The claim that APSS/PI accurately predicts objective performance is completely unsubstantiated by the data offered. An alternative analysis should have separated screeners into groups based on performance (errors), not APSS/PI scores. A "not recommended" group could have represented all screeners with two or more errors while a "recommended" group could have been composed of all screeners with no screening errors. Comparison of these groups with regard to APSS/PI scores would have generated more interpretable results.

In fact, using the APSS/PI to make hiring decisions would have eliminated 40 percent of those not recommended for hire who had error free performance while employing a third of the recommended sample who had performance errors. The relationship between the overall negative performance composite score and APSS/PI ($r = +.16$) indicates little improvement in prediction of satisfactory performance by incorporating job tardiness as a measure of performance. Overall, the data do not justify use of the APSS/PI in a selection environment. The selection algorithm clearly has an unusually high false positive rate and is based on a weak relationship using very small sample sizes.

Perhaps the most confusing aspect of the data is the relationship between supervisory ratings and the APSS/PI. The APSS/PI does not predict supervisory performance evaluations. The investigators simply explained that supervisory perceptions are not an accurate reflection of objective job performance. While this may be true, no data are presented that examine the relationship between testing errors and supervisor evaluations of performance although these data were available to conduct such an analysis. An alternative explanation could be that supervisory evaluations were an accurate prediction of objective job performance but the APSS/PI is not sensitive enough to predict these differences. It is never clear why the contractor completely rejected the use of the subjective ratings to validate the screener's job performance.

A final issue is with regard to the effect of experience on job performance. Although the data were available, no comparison was made between months on the job and any testing errors measures. In a selection and training system it is important to understand the effects of experience or training on actual job performance. It is plausible that after several months of experience as a screener, performance would improve such that errors in target detection become an unlikely occurrence. The contractor studies never examined experience as a possible factor.

While these investigations demonstrated the need and requirement for effective personnel selection practices, they do not offer a valid, reliable, and comprehensive instrument that can be used across the aviation security industry. Perhaps the industry's overall rejection of implementing this tool is the best indication of its inadequacies.

Although the contractor was the only research-based effort that attempted to identify unsuccessful screeners prior to employment by using a personality trait model, other efforts have sought to distinguish successful from unsuccessful airline screeners using demographic characteristics. However, current statutes regarding hiring practices have negated the use of much of this work in a selection environment.

3.3.2.2 International Total Services (ITS) Study.

One aviation security contractor recently sought to determine if there is a relationship between screener performance and specific employee characteristics. In 1993, the Training and Personnel Development Department at International Total Services (ITS) completed a study of demographic factors and their relationship to job longevity and performance on FAA checkpoint testing (ITS, 1993). ITS compared these performance criteria for a sample of 3,183 screeners divided into sub-groups by: (a) age, (b) sex, (c) ethnic background, (d) educational level, (e) previous employment, (f) military background, and (g) citizenship. Although these results can not be used for employee selection and hiring, some findings are nevertheless worthy of mentioning.

In a comparison of screeners with three or more years tenure versus the total sample, several significant differences were identified. When compared to the nationwide screener population, high tenured screeners were on the average, older, better educated, more likely to be female, more likely from Asian than black ethnic background, and less likely to possess U.S. citizenship (see table 1). Nearly one-third of the high tenured sub-group were of Asian background (nearly equally divided by gender), with another 17 percent Caucasian females. Employment tenure was also found to be twice the duration for the "housewife" subgroup than any other prior work experience sub-group. Perhaps the most surprising finding however is the relationship between educational level and job tenure. ITS reports that screeners with four year college degrees and those with eight years of college have the two highest employment duration averages, respectively. This probably reflects the current poor outlook in the job market for college graduates and a work environment that is attractive for retired professionals. Interestingly, demographic factors were not useful in differentiating screeners who remained employed less than 60 days from the total sample.

TABLE 1. COMPARISON OF SCREENERS NATIONWIDE TO VETERAN SCREENERS
(DATA FROM ITS PERSONNEL PERFORMANCE STUDY, 1993, SECTION III:
EMPLOYEE RETENTION, GRAPH 1)

	TOTAL SAMPLE OF SCREENERS (N = 3,183)	SCREENERS WITH ≥ 3 YEARS WORK EXPERIENCE (N = *)
Average age	38 years	54 years
Female	45.4 percent	57.2 percent
Asian	10.9 percent	30.9 percent
Black	32.4 percent	12.2 percent
U.S. Citizen	82.9 percent	63.0 percent

* The sample size of the veteran screeners was not reported.

Of potentially great importance is the ability to identify screeners who do not perform effectively (i.e., fail FAA checkpoint tests). The data presented by the ITS study however, do not demonstrate a relationship between demographic factors and the number of FAA test failures over the preceding 12-month period. The only substantial difference found is that the black subgroup is more likely to fail FAA tests than any other ethnic group.

Similar to findings for the job tenure criteria, several demographic factors distinguish superior performance in screeners (i.e., as determined by security firm evaluations) from the total nationwide sample. A group of 159 screeners were selected for special recognition because they "... showed an exceptional ability to repeatedly pass FAA tests." This selected group differed from the overall nationwide sample in: (a) mean age (51 versus 38 years); (b) gender (54 percent versus 45 percent female); 3) ethnic background (50 percent versus 24.2 percent Caucasian); (d) FAA pre-employment average test scores (98 versus 93.5); and (e) experience background (nearly a three-fold difference in prior military background and twice the percentage of the total sample with a law enforcement background).

As previously stated, the use of these data in making hiring decisions is unacceptable under numerous federal and state statutes. From a methodological standpoint, there are a number of critical issues regarding the data that need to be addressed that impact the interpretation of the findings as well.

All the ITS data are presented in actuarial format using frequency tabulations as the primary means for the analysis. There were no statistical methods applied to the data to assess either the magnitude of the differences shown, or to determine if statistical significance existed. Non-parametric techniques (e.g., Chi-Square analyses) could have been used on nearly all the data and would have provided a more meaningful interpretation.

A second critical issue concerns the methods used to sub-divide the sample population. In most cases, the data are simply grouped along one dimension (e.g., gender) rather than multiple dimensions (e.g., gender X age X ethnic background). Consequently, assignment to a subgroup is severely confounded. For example, the sub-groupings within the age grouping can be

confounded by educational level by a high proportion of retired professionals. Similarly, citizenship assignment could conceivably be compromised by one specific ethnic group. Without considering the multi-dimensional nature of the demographic factors involved, it is difficult to isolate the effect of specific factors.

Most importantly, the data are only partially useful in identifying superior performers (i.e., as assessed by company evaluations) or job longevity post facto. Regardless, these factors cannot be used for promotion decisions or incorporated into employee recognition programs. Although not conclusive, the findings offer no evidence for predicting which screeners have a propensity for committing screening errors. In addition, the results are not useful in identifying the new applicant who will leave the career field within the first two months. A strong selection system must have the potential of identifying the short-term or ineffective applicant before employment.

3.4 AIRLINE PASSENGER SECURITY SCREENER TRAINING.

A well-trained staff is key to providing an effective airline security program. Most of the emphasis however, has been placed on developing technology (e.g., neutron techniques, gamma ray techniques) and very little has been directed toward enhancing human capabilities.

Most screeners are trained with a limited inventory of obvious test items and are only taught to recognize complete items—not their individual components. Some security firms though, were found to exceed the required training minimums. These firms have developed their own cirriculum to supplement the FAA/airline requirements as well.

3.4.1 Syllabus Topics.

Currently, most security companies provide a three-part training program for airline security screeners. Part one incorporates 12 hours of classroom training over a three day period (see table 2. Part two involves 40 hours or supervised on-job-training (OJT), while part three entails recurrent training scheduled after a six-month period.

In order to satisfy the classroom training requirements, the ATA has developed the "Airline Security Screener Training" manual, an instructor's guide approved by the FAA for use in checkpoint security training. The manual includes 13 lesson plans in order to provide guidance to the instructors and ensure a comprehensive course of instruction. Table 2 below briefly illustrates each lesson and the allotted time.

TABLE 2. ATA-APPROVED CLASSROOM SYLLABUS FOR CHECKPOINT SECURITY SCREENERS

LESSON PLAN	TOPIC(S)	DURATION (MINUTES)
Session 1	The First Line of Defense (video)	10
	The Reason for Passenger Screening	10
	The Legal Basis for Passenger Screening	15
	The Effectiveness of Passenger Screening	15
Session 2	How Aviation Security Works	50
	Screening Techniques (video)	
Session 3	The Importance of the Screener	50
	Screening Techniques (video)	
Session 4	Identifying the Threat	50
	Screening Techniques (video)	
Session 5	Identifying the Threat	50
Session 6	Screening the Passenger	50
	Screening Techniques (video)	
Session 7	Screening the Passenger	50
	Screening Techniques (video)	
Session 8	Screening Hand-Carried Items	50
	Screening Techniques (video)	
Session 9	Screening Hand-Carried Items	50
	Screening Techniques (video)	
	X-Ray Image Interpretation Teaching Aid (video)	
Session 10	Screening Atypical Passengers	50
	Screening Techniques (video)	
Session 11	Special Situations	50
Session 12	Security Information Circulars and Security Directives	25
	Denied Entry into the Sterile Area	25
Session 13	Guidelines for Station Specific Instruction	as necessary
	It's Your Responsibility (video)	

The Transport Canada's Security Training and Development Division has also developed a "Screening Personnel Course" training manual for security screeners in Canadian airports. The manual is currently being revised to include new training procedures and policies. To compare the two training programs, the following topics from Transport Canada's manual are illustrated below:

a.   Introduction to Screening (job familiarization, duties, responsibilities, and authority)·
    ?

b. Weapons, Explosives, Dangerous Articles and Dangerous Substances (types/kinds of threats, threat identification, how to react, who to contact);

c. The Walk Through Metal Detector (types of metal detection units, operational features, functional guidelines);

d. The Hand Held Metal Detector (purpose, familiarization, limitations, operational procedures, maintenance);

e. Body Search of Persons (purpose, applicability, guidelines and procedures);

f. Scanray Linescan® System 1 and 2 (familiarization, safety precautions, operational procedures, law enforcement assistance);

g. Hand Search of Baggage (when required, baggage requiring special consideration, exceptions, guidelines and procedures);

h. Special Situations (diplomatic missions, aircraft crews, handicapped persons, persons escorting prisoners, etc.);

i. Terrorism (types of terrorism, terrorist incidents, terrorist profile and common characteristics).

The OJT phase (part two) is typically not formalized, with little standardization between companies. However, the OJT environment is really where screeners develop the critical expertise necessary to perform their job successfully. There is even less emphasis placed upon the importance of recurrency training (part three) in many companies. This phase of training is often based upon a review of the ATA training tapes illustrated in the beginning of this section. In consideration of the high turnover rates in this industry, it is obvious that a significant percentage of screeners never reach this final phase of training.

3.4.2   Research on Training to Improve Screener Performance.

An early study (Potter, 1975), as reported in Funke (1979), examined the ability of perceptual training techniques on developing the skill of detecting weapons for screeners. Data were obtained from 18 subjects, of which one-half were college students. All were high school graduates. Subjects were randomly assigned to one of three groups: Group 1 was trained with sound/slide presentations dealing with explosives; Group 2 was trained with an audio/visual program on guns; and Group 3 (the control group) was trained with written instructions concerning items they should watch for during visual inspections. Subjects were shown 95 slides of X-rayed baggage of which approximately 30 percent contained various weapons. Each slide was shown for 6 seconds followed by a 4 second blank period. Subjects were to indicate whether they thought the bag required opening, and if so, the reason for their decision.

19

Three measures of performance were analyzed: (a) number of bags containing weapons which were not found (i.e., missed detections); (b) number of bags which were incorrectly identified as containing weapons (i.e., false alarms); and (c) number of bags correctly identified as containing weapons but the weapons themselves misidentified. The number of false alarms and the number of misidentified weapons were not significantly related to the training conditions. However, the type of training had a significant effect on the number of weapons found, although no statistical data were reported by Funke (1979). The mean percentage of weapons missed for the three groups were reported as follows: explosive trained ($\underline{M}$ = 10.0), gun trained ($\underline{M}$ = 22.6), and control group ($\underline{M}$ = 32.8). Funke indicated that these findings suggest that perceptual training techniques (e.g., sound/slide presentations) can significantly improve target identification on visual inspection tasks, particularly when targets appear infrequently.

Another study by Nadler and Mengert (1993), investigated the benefits of computer-based instruction (CBI) compared to current methods of selection, training, and screener certification. They used the Safe Passage™ System, a CBI system developed specifically for airline passenger security screeners. The Safe Passage™ System presents screeners with X-rayed baggage images stored in a video database of approximately 2,000 images. The images represent eight categories: (a) innocent, (b) suspicious innocent, (c) electronic innocent, (d) explosive, (e) gun, (f) knife, (g) other sharp objects, and (h) combined/other weapons.

Nadler and Mengert (1993) were given data on 1,465 screeners who worked for security companies at five major domestic airports equipped with the Safe Passage™ System. The preliminary findings, as reported in their interim report, are based on data from 500 screeners referred to as their Sampled Data Set. Screeners voluntarily performed "simulated" screening tests while threat images were presented randomly. Each test contained 12 different images. Three levels of proficiency were programmed (i.e., low, medium, high) so that each level contained increasing percentages of relatively "difficult" test images: 25, 50, and 75 percent, respectively.

They found that four of the eight image categories (i.e., suspicious innocent, explosive, knife, other sharp object) showed accuracy less than 80 percent in the Low Proficiency Level, reporting accuracy levels of 75.1 percent, 77.4 percent, 76.8 percent, and 77.3 percent, respectively. Thus, images that fall into these categories can generally be more difficult to detect. Preliminary results also indicate a low percentage of "critical errors" (i.e., errors resulting when a screener passes a bag that should have been held for further inspection). They noted that 50 percent committed no critical errors at each proficiency level. Based upon these initial data, Nadler and Mengert (1993) concluded that there is, however, ample room for improvement in screener's ability to distinguish threat items from innocent items in X-ray images of baggage. This CBI system appears to be the most effective and advanced training program available on the market today.

In the most recent study to date on airline screeners, Kaempf, Klinger, and Wolf (1994) employed a cognitive systems engineering approach (CSE) to identify the decision requirements of security personnel using X-ray scanning devices at airport security checkpoints. Their primary objective was to develop decision-centered interventions that will enhance screener performance

either through training or through the implementation of decision supports (e.g., pattern recognition system). Forty-three security personnel were interviewed, including managers, trainers, screeners-in-charge, and checkpoint screeners.

"Screening companies must recruit individuals who have no applicable skills for minimum wages and instill some measure of proficiency in a relatively short period of time" (Kaempf et al., 1994, p. 29). They also identified several serious problem areas with current training programs. First, none of the existing training programs address the cognitive skills required for screeners to perform their jobs proficiently. Second, the OJT program is not a well-structured, controlled effort designed to enhance the skills of the novice screener. Finally, recurrent training is no more than a review of training received during initial hiring and is similarly ineffective.

Phase I of this study (Kaempf et al., 1994) has yielded information highlighting the differences between novice and expert screeners in their decision-making processes. Their findings indicated that experts have developed, through accrued experience, an extensive mental "library" of both non-threat and target (threat) items. This mental "library" facilitates the screener's ability to rapidly and accurately differentiate between the preponderance of non-threat items and the infrequent threat items—a critical skill. Kaempf et al. stated, "What is needed is an integrated training program that considers the skills required to perform proficiently and that uses training opportunities to build on skills previously developed" (p. 30).

3.5   SCREENER TURNOVER RATES.

Most aviation security firms experience high turnover rates—in some cases, 100 percent in just 10 months. In their recent study on airline screeners, Kaempf et al. (1994) identified several factors which lead to such high turnover of personnel, including low pay, lack of opportunities for advancement, and competition for these types of workers in the job market. Kaempf et al. stated, "The facilities that we visited typically have a stable employee base of 10 percent with the remainder in constant flux. This results in about 90 percent of all screeners at any given checkpoint having less than six months experience" (p. 29). High turnover rates such as these will ensure that very few screeners will become highly experienced.

The extremely high turnover rates so prevalent in the industry for the airline security screener position have two critical implications. First, it inhibits the development of expertise; and second, it reduces the ability to take advantage of an effective and cost-efficient method of training (i.e., OJT from experienced screener "mentors"). "The constant migration of personnel creates significant requirements for training that strain available resources. For example, there is a continuous need to conduct initial training for new hires" (Kaempf et al., 1994, p. 29).

An examination of the starting wages offered to new hires presents some possible evidence for such high turnover rates (see table 3). As a general rule, occupations with such high responsibility normally necessitate higher levels of compensation and benefits.

21

TABLE 3. STARTING WAGES AND BENEFITS FOR AIRLINE PASSENGER SECURITY SCREENERS BY AIRPORT AND SECURITY CONTRACTOR.

BENEFITS	LOCATION	COMPANY	HOURLY WAGE
	JFK Airport	Argenbright	$4.50
Yes			
	Newark Airport	ITS	$5.05
No			
	O'Hare Airport	Andy Frein	$4.25
No			
	Orlando Airport	Argenbright	$4.25
No			

3.5.1   Incentives/Reward Systems.

Several years ago, Hertzberg (1966) noted that achievement, recognition, and responsibility are the most consistent motivational factors producing job satisfaction. Many theories state that good job performance leads to job satisfaction rather than job satisfaction leading to good job performance (Brennan, 1974). Wallis (1992) reported that reward systems generally stimulate attention and have proven effective in many airport security checkpoints. An effective incentive program can also lead to significant gains in productivity.

Wiener (1975) argues that an offer of a 'substantial financial reward' does not significantly influence accuracy in inspection tasks. According to Wiener (1975), Murrell (1965) explains, "It must be realized that inspection is largely a perceptual task and ... successful perception is not a process which can be readily influenced by an incentive, however great" (p. 110). Furthermore, Wiener (1975) noted that financial incentives in an experimental setting have generally yielded unimpressive results in inspection studies.

3.6   VIGILANCE OPERATIONAL SIGNIFICANCE FOR AIRLINE PASSENGER SECURITY SCREENERS.

For nearly 40 years, vigilance researchers have engaged in one of the most prolific programs of scientific research in any area of behavioral science. Well over a thousand studies have been published during this time (Mackie, 1987). Over 30 years ago, Buckner and McGrath (1963) conducted a symposium on vigilance and addressed the following questions: "Why do people fail to detect important signals? Under what circumstances do they fail to detect them? What can we do to ensure that they detect them?" (p. vii). This section of the report will attempt to explore these questions as they relate to the airline passenger security screener occupation.

Human vigilance is the ability to detect rare signals over a prolonged period of time. Simply put, it is the ability to maintain attention. Vigilance has been described in several ways: as performance on monitoring tasks; as attention over extended periods of time; as a state of

maximum physiological efficiency; and as a state of readiness to respond to infrequent, yet small changes occurring at unpredictable random intervals (Buckner and McGrath, 1963; Davies, 1979; Macworth, 1957). Vigilance is synonymous with long-term attentive behavior (Davies and Parasuraman, 1982) and sustained attention (Parasuraman, 1986). According to Jerison and Wing (1963), vigilance is a general problem of attention or alertness. The word attention however, unlike vigilance, is used with little confusion in daily affairs (Edwards, 1990). Other terms used in the literature include monitoring and search—all are treated as aspects of attention (Parasuraman, 1986). The authors will not make any distinction between vigilance, attention, monitoring or search in this report. The reader is encouraged to review Parasuraman (1986) for complete definitions and key concepts and terms.

Vigilance performance is required in a number of industrial settings. To perform successfully, operators must remain alert for indications of malfunctions and changes in operating states. For example, the air traffic controller who keeps aircraft separated by observing 'blips' on a monitor; the flight engineer who observes various engine performance parameters (e.g., oil temperature, oil pressure, fuel flow) to ensure a safe flight; the quality control inspector who examines a number of products (e.g., circuit boards, microchips, beverages) to detect and remove deflective or flawed items; and the anesthesiologist, who monitors a patient during an operation to ensure all vitals signs are within an acceptable range (Howland and Wiener, 1963; Wickens, 1992). Our defense against a nuclear attack depends ultimately upon the vigilance of the persons observing the displays in our early-warning stations throughout the world (Buckner and McGrath, 1963).

3.6.1 Origins of Vigilance Research.

Research on vigilance, monitoring, and search arose in response to problems identified during World War II. In 1943, the Royal Air Force (RAF) asked if experiments could be conducted to determine the optimum length of watch for radar operators on anti-submarine patrol. RAF reports documented that a number of potential U-boat contacts were being missed and that the operators were overstrained (Davies, 1979; Davies and Parasuraman, 1982; Parasuraman, 1986). After some preliminary experiments conducted by Mackworth (1950), the RAF began an operational study of the detection of submarines by radar operators. They found that after 30 minutes on watch, a marked deterioration in performance occurred—thus coined the term "vigilance decrement." This performance decrement has been a recurrent finding in countless laboratory studies observing human behavior and will be discussed in a subsequent section of this report.

3.6.2 Theories of Vigilance.

There have been a number of theories advanced over the years that attempt to explain the vigilance phenomenon (e.g., arousal theory, expectancy theory, inhibition theory, signal detection theory). Although no particular theory is likely to be entirely correct, Wickens (1992) noted that the advantage of such theories is that they provide ways of accounting for vigilance loss and thereby suggest techniques to improve vigilance performance.

Two of the more popular vigilance theories that have applications for airline passenger security screeners are the *expectancy theory* and the *filter theory*. According to Davies and Parasuraman (1982), the expectancy theory hypothesizes that the observer develops expectancies about the future probability of a signal occurrence on the basis of previous experience(s) with the task, and that the detection rate is determined by an expectancy level. For example, if the probability of occurrence is low, as it is for screeners, then the expectancy for observing a signal is low. Based on this theory, Wickens (1992) noted that introducing false signals will help keep the screener attentive, particularly if the signals are physically similar to the real signals themselves. Thus, according to the expectancy theory, the periodic tests carried out by the FAA and airlines not only allow for a measure of performance, but also helps keep screeners on their best possible vigilant behavior.

For the airline passenger security screener to maintain a 100 percent detection efficiency using the expectancy theory, the screener must assume that every bag and every passenger contain or possess a weapon of some sort. Although realistically, 100 percent detection efficiency seems virtually impossible, some simple techniques have shown signs of progress. For example, Wickens (1992) reported that simple instructions describing new policies and/or procedures can dramatically improve detection performance. Wickens stated, "... in airlines security inspection, increased stress on the seriousness of misses (failing to detect a weapon smuggled through the inspection line) caused a substantial decrease in the number of misses from 1987 to 1988" (p. 46). The source(s) of these data however, were not reported by Wickens.

The second theory that has implications for the airport screener is the *filter theory*. Developed by Broadbent in the 1950s, the filter theory is based on experiments conducted in the areas of selective attention, vigilance, and the effects of noise on performance (Davies and Parasuraman, 1982). The filter theory proposes that there are periodic failures on the part of the operator to take in task-relevant information and the net result is a decrease in the number of detections (i.e., missed targets) and an increase in reaction times (i.e., longer time to react). Loeb and Alluisi (1984) note that this performance decrement is particularly marked when operators are monitoring channels of information where there is considerable repetition (e.g., screening for weapons). According to Davies and Tune (1970), the temporary failure to take in task-relevant information is attributed to the presence of competing stimuli, since different classes of stimuli (e.g., size, shape, complexity, novelty) are competing with one another to enter the filter. The filter can however, be trained to select certain types of information (e.g., guns, knives, dynamite sticks) by properly instructing the operator through repeated practice or test sessions.

The filter theory assumes that the operator's information-handling capacity is limited and that he/she cannot analyze more than one stimulus simultaneously, therefore competing stimuli are selected sequentially for analysis (Loeb and Alluisi, 1984). The selection of each stimulus and the order in which they are analyzed depends on certain stimulus features, stimulus importance, and novelty of the stimulus (Broadbent, 1958).

3.6.3 Vigilance Decrement.

In some vigilance situations, although not all, the number of signals correctly detected (i.e., detection or hit rate) has been found to decline with time on task (Davies and Tune, 1970). Fox (1975) reported that with prolonged inspection periods, detection performance can deteriorate rapidly, and "... drops of 40 percent in 30 minutes have been noted" (p. 89). The magnitude of the vigilance decrement in various studies has been inconsistent however, with some studies reporting miss rates as high as 40 percent (Craig, 1984).

Since the early efforts, research on vigilance has primarily been concerned with reaching an understanding of the factors which are responsible for this decrement in performance (Davies and Parasuraman, 1982). The vigilance decrement is simply the inability to maintain vigilant behavior (i.e., attention or alertness) over extended periods of time (Pierce et al., 1991). This performance characteristic has been a focal point of the "critics challenge" and is generally uncritically accepted, even though some investigators (Smith and Lucaccini, 1969; Wiener, 1975; Wylie, Mackie, and Smith, 1985) seriously doubt its existence in real-world settings. They contend that the vigilance decrement only occurs in laboratory settings and not in the real world, and is merely a result of fading motivation of the subjects. Others (Nachreiner, 1977; Wiener, 1987) support the same viewpoint. Wiener (1987) stated: "There is no evidence ... to suggest that vigilance decrements exist in real-world systems and, indeed that it is anything but a laboratory artifact brought on by contrived experimental tasks, unmotivated subjects, and poor instructions" (p. 730). Furthermore, Davies and Tune (1970) noted, "... under such conditions [i.e., laboratory experiments], it is not surprising that observers sometimes fall asleep and that reports of drowsiness and extreme boredom are common" (p. 11).

There has not been an experiment to date that either confirms or denies the existence of the vigilance decrement (Wiener, 1987). While there are ample studies suggesting that the vigilance decrement does in fact exist, some field studies have yet to encounter this problem. For example, Nachreiner (1977) reported on the performance of coin inspectors over several days and they did not reveal any decrement with time on task.

There have been a number of theories advanced that attempt to explain the vigilance decrement (e.g., arousal theory, habituation theory, motivation theory). Two of the more popular theories are the *arousal theory* (or activation theory) and the *motivation theory*. According to Davies and Parasuraman (1982), the *arousal theory* maintains that a progressive reduction in the arousal level of the central nervous system takes place during task performance. This decrease in brain activity is largely brought about by the monotonous nature of the vigilance situation, and as a result, the brain becomes less responsive to and less efficient at dealing with external stimulation. On the other hand, the *motivation theory* maintains that the vigilance decrement is attributable to individual differences in motivational level (e.g., some people being more conscientious monitors than others), as well as to reductions in motivational level caused by the monotonous conditions of work and the failure to provide adequate incentives. It is frequently viewed as being of crucial importance in the debate about the practical significance of vigilance research (Davies and Parasuraman).

Other investigators (Pierce et al., 1991) purport that the vigilance decrement depends upon the task-related variables (e.g., frequency or density of monitoring, display size, display brightness, viewing angle, viewing distance) and the individual characteristics (e.g., age, intelligence, motivation, personality). More on how these characteristics can affect vigilance performance is reported in a subsequent section of this report.

Nachreiner (1977) presented a different viewpoint regarding the potential reasons for this apparent performance decrement. Based on his findings from an optical vigilance experiment, Nachreiner reported that performance decrements in vigilance experiments are dependent on the subject's perception and realization of the experimental situation. Nachreiner recruited subjects by two announcements: one asking for participation in a psychology experiment, and the other asking for applicants for the job of experimenter. Eight subjects were equally divided into two groups, the control group and the experiment group. The subjects had to distinguish a series of spikes simulating bioelectrical potentials where the critical signal had 10 spikes and the neutral signal had eight. Nachreiner found a decrement after the first 30 minutes for the experiment group, but not for the control group. These findings suggest that the vigilance decrement might be more dependent on the subject's perception of the experimental condition rather than on the specific task characteristics of the experiment conducted. These findings however, should be viewed as inconclusive because of the limited number of subjects.

3.6.4   Vigilance Tasks.

Most vigilance tasks entail observing a display and monitoring it for very subtle or obvious changes in state. Davies (1979) stated that vigilance tasks are tasks "... in which attention is directed to one information display ... over long, unbroken, periods of time, to detect infrequent changes in the state of the display that are extremely difficult to discriminate" (p. 14). Other investigators (Buckner and McGrath, 1963; Davies and Parasuraman, 1982; Edwards, 1990) refer to vigilance tasks as "monitoring" or "watch keeping" tasks.

Craig and Colquhoun (1977) stated that a major criticism surrounding vigilance research is that most tasks used in laboratory studies are too simple to have any relevance for real-life operations. Nachreiner (1977) stated, "There are many differences between vigilance experiments in the laboratory and monitoring jobs in field situations. Because of these differences, the question has been raised whether vigilance experiments are relevant to the problems of monitoring in field situations" (p. 666). Wiener (1987) stated: "... the nature of the monitoring task, which operates in a real-time, unpredictable, event-driven environment for long periods day after day, it is unlikely that any laboratory experiment could be set up even to faithfully simulate, let alone duplicate, a working system" (p. 727).

Numerous investigators have examined the effects of vigilance on simple and complex tasks. However, most studies were difficult to compare because they employed different kinds of vigilance tasks, different response types, and different response measures (e.g., hit rate, reaction time, false alarms). Wiener (1987) stated: "The 'monitor' actually time-shares between monitoring and more active tasks [e.g., searching suspicious bags], and thus the job is not as passive, routine, or cognitively unenlightening as the papers on vigilance would imply" (p. 730).

26

Mackie (1987) stated: "Countless variables have been identified that may influence the decrement or level of performance" (p. 707).

Table 4 presents 15 of these variables that influence vigilance performance. Many of these reported variables do not represent an immediate threat to airline passenger security screeners. For example, performance decrements associated with *task duration* are often countered by requiring screeners to rotate positions every 20-30 minutes. This rotation reduces the probability of error(s) associated with the task monotony. In addition, improving employee recognition would work as an incentive by boosting motivation and morale among employees. Implementation of such a program could be quite successful, as well as very cost-effective. Other factors however, (e.g., circadian rhythms, payoff and/or rewards, motivation, morale, social environment), present some very obvious challenges to the screener occupation. Very little is known how these factors would impact screener performance; thus, further research is needed to address these aspects of screener tasks.

TABLE 4. PARTIAL LISTING OF FACTORS THAT AFFECT PERFORMANCE IN VIGILANCE TASKS (MODIFIED FROM MACKIE, 1987, P. 708)

Task duration
Work/rest pattern
Circadian rhythms (first shift versus third shift)
Sleep loss; sleep quality
Incentives; payoff and/or rewards
Motivation (intrinsic, extrinsic)
Display design (rate of presentation, display size, lighting/illumination)
Individual characteristics (age, gender, intelligence, experience, intro-extrovert)
Mood or morale
Environmental stressors (noise, sound, vibration, lighting, temperature)
Social environment (peer pressure, supervisor/management pressure)
Illness; injury
Drug use (caffeine, nicotine)
Exercise, physical work

3.6.5   Performance Measures in Vigilance-Related Tasks.

Performance measures taken from vigilance tasks are usually a measure of operator efficiency. Most measures are generally assessed by one of three ways: (a) the number of correct detections (known as the *detection rate* or *hit rate*); (b) when a signal has been reported and none has been presented (known as *false alarms* or *false detections*); and 3) when a signal has been presented but not identified (known as *missed detections*). A fourth measure, but not as common, is the amount of time taken to detect a signal, referred to as the *detection latency* (Davies and Parasuraman, 1982; Davies and Tune, 1970).

The most commonly used measure of performance in vigilance studies is the detection rate (Davies and Tune, 1970); however, much useful information can be obtained from the other performance measures. According to Davies and Parasuraman (1982), the primary measures of performance in vigilance situations are the detection rate, the false alarm rate, and the detection latency. All three measures are essential for understanding the way in which vigilance performance varies with time on task, across different experimental conditions, and between different individuals.

In an experimental setting, performance measures for airline screeners could include: (a) the detection rate (i.e., number of bags correctly identified as containing weapons); (b) false alarms (i.e., number of bags incorrectly identified as containing weapons); and 3) missed detections (i.e., the number of bags containing weapons which were not found). This method was employed by Potter (1975) as previously reported on page 31.

In industry settings however, measuring vigilance performance is difficult to say the least. For example, the number of missed detections (i.e., weapons that make it past the screening checkpoint) is not known; therefore, it is difficult to assess this particular aspect of performance. Most of the tests used to measure vigilance performance in industry settings are validated against supervisors' ratings (Wiener, 1975). This subjective criteria is questionable however, since the supervisor may be totally unaware of the actual performance of the individual inspectors. In light of this issue, Wiener stated, "His ratings [the supervisor's] are probably based on perceptions of earnestness and cooperation, and the correlation between these attributes and actual inspection performance is unknown" (p. 102).

Despite the inherent limitations in measuring vigilance performance in the actual working environment, by using test weapons such as those currently employed by the FAA and air carriers, we could at least identify those screeners who fall below the accepted standards. Once identified, these screeners could then be sent for additional training to focus on their area(s) of weakness, thus theoretically, improving the overall efficiency of the security system. Nonetheless, the FAA and airline "tests" lack the ability to accurately determine who the better performers are because of their apparent and repeated use of similar test objects.

3.6.6 Techniques to Combat the Vigilance Decrement.

There are some methods and/or techniques that can increase vigilance performance. Wickens (1992) noted that any technique that will enhance the subject's memory of the signal characteristics (e.g., object shape, size, density) should reduce performance decrements and preserve a higher level of awareness. For example, by introducing false signals (e.g., test weapons), one improves or heightens the operator's awareness, consequently improving the system's effectiveness. Furthermore, Kelly (1955) as reported in Wickens, found a large increase in detection performance when quality control operators were allowed to look at television pictures of idealized target stimuli. This emphasizes the importance and continued practice of using FAA and airline test objects to continually refresh the screener's memory and heighten their awareness, despite the objects' obvious shortcomings.

Another method to combat the vigilance decrement was proposed by Childs (1976). He recommended that inspectors should have access to visual representations of possible defectives rather than the representation of those that are normal. This position lends further support to the expansion of the FAA and airline test items to include "parts" or "components" of a threat, rather than the complete and assembled item. Other investigators (Davies and Parasuraman, 1982; Craig, 1984) have also suggested different methods for reducing the vigilance decrement. Most suggestions are extrapolations from laboratory experiments; they include:

    a.   provide appropriate work/rest cycles;

    b.   instill motivation by emphasizing task importance;

    c.   introduce mild environmental stress (heat, noise, vibration);

    d.   provide observation by supervisors;

    e.   use personnel selection techniques to identify individuals with a propensity for maintaining vigilance.

3.7   SELECTION TEST DEVELOPMENT FOR U.S. ARMY PERSONNEL.

Very few laboratory studies have involved watch durations or performance tasks that even remotely resemble airline passenger security screeners. However, of particular interest is the recent work of Crumley, Pierce, Schwalm, Coke, and Brown (1992). They conducted an experiment with U.S. Army enlisted personnel to determine if cognitive style tests could be used to predict target detection performance. Such tests, if validated, could prove invaluable to the airport security industry while screening for good performers during the pre-employment process. The purpose of their research was to determine if scores from a group of cognitive factors (i.e., Speed of Closure, Flexibility of Closure, Perceptual Speed) and tests (i.e., Gestalt Completion, Concealed Words, Snowy Pictures, Hidden Figures, Hidden Patterns, Identical Pictures), or selected biographic items (i.e., gender, smoking habits, coffee drinking) predict the ability of persons to detect targets. They reported that the Speed of Closure and the Flexibility of Closure factors were used because they appear to "... involve factors important to the task of seeing partially obscured items" (p. 6). They further noted that the Gestalt Completion Test was used because its items appear to be almost identical to the visual and intellectual tasks involved in detecting camouflaged targets where portions of the items blend into the background.

Crumley et al. (1992) presented 1,440 slides developed from photos taken with a 35mm camera from a helicopter flying between 400 and 800 feet over and near Fort Still, Oklahoma. The slides depicted aerial views of terrain where subjects were to identify man-made objects other than roads, railroads, power lines, and fences (e.g., vehicles, bridges). Of those 1440 slides, 90 (6.3 percent) had actual targets. The slides were presented on a 19" black-and-white television screen. Data from 209 subjects were used in the analysis. Two measures of performance were used in their experiment: (a) targets detected, and (b) false detections.

Their findings indicate that certain biographical items can be found to affect a person's ability to detect targets. They found that persons who fish for amusement make more target detections than those who do not fish (F = 3.30, p = .039). Their rationale for this finding is that fishing is often appealing to persons who do not become restless in situations where attention must be maintained over prolonged periods. They also found that people who do not wear glasses make more detections than people who have glasses prescribed (F = 5.09, p = .025). They noted that this was probably a result of glasses not being worn when they should have been. Lastly, they found that persons who drink moderate or small amounts of coffee daily (i.e., less than 6 cups per day) make more detections than heavy coffee drinkers (i.e., those who drink more than 5 cups per day) (F = 4.34, p = .014). Stimulants (e.g., amphetamines, caffeine) in small to moderate doses can improve detection performance by increasing the arousal level of subjects. This finding has also been reported by other investigators (Davies and Parasuraman, 1982; Mackworth, 1965; Wickens, 1992).

Based on their findings, Crumley et al. (1992) concluded that the Snowy Pictures test is perhaps the best test for predicting target detection performance, and that it can be used in research situations where it is desirable to identify subject aptitude before the subjects are tested. Furthermore, they noted that the cognitive factor Speed of Closure and possibly the Flexibility of Closure factor also predict target detection performance. They suggested that cognitive tests could form a basis for selecting persons needing additional training and development after they are assigned to their particular occupational specialty. These findings appear relevant to airline passenger security screeners because of the task similarities (i.e., searching for and detecting hidden targets), however warrant further investigation.

3.8   EFFECTS OF STRESSORS ON SONAR OPERATORS.

Sonar operators play a critical role to military missions of the surface, sea, and air. The tasks involved in sonar operations are similar to those of airline passenger screening. For example, sonar operators must watch a display for extended period of time looking for small and subtle changes in the display state. Wylie, Mackie, and Smith (1985) surveyed 212 Royal Navy sonar operators (i.e., submarine operators, surface ship operators, helicopter operators) to gain a better understanding of the stress conditions associated with the occupation and to identify countermeasures that will help minimize those stress-related effects.

Most stressors are assumed to have a negative impact on operator performance, although the degree of impact or degradation is unclear (Wylie et al., 1985). This uncertainty prompted Wylie et al. to collect data on the perceived impact and prevalence of stressors associated with sonar operations. They divided stressors into four separate categories: (a) task-related stressors (boredom/monotony, operator workload, displays and controls, work station/personnel equipment design, midnight and morning watches); (b) environmental stressors (uncomfortable heat, uncomfortable cold, vibration, noise, uncomfortable air pressure, illumination or lighting problems); (c) organismic stressors (fatigue, tiredness, minor illness); and (d) social stressors (command pressure). Note that no statistical data were presented in their report.

They found that there was a strong agreement among sonar operators concerning which stressors had the most serious impact on operator performance. The two most common stressors reported were boredom and fatigue. Boredom, generally associated with monotonous operating conditions, was viewed as the "worst" stressor impacting performance. The adverse impact(s) of boredom is clearly evident in most monitoring occupations. The respondents reported that boredom/monotony occurred very frequently (in fact, more so than any other stressor); therefore, it was perceived as more severe than any other stressor. Other investigators (Boff and Lincoln, 1988; Thackray, Bailey, and Touchstone; 1977a) found similar detrimental effects of boredom (i.e., subjects that give self-reports of high boredom typically show longer reaction times and greater performance decrements).

The impact of fatigue on operator effectiveness was also reported as a recurring problem. Fatigue was ranked very high (i.e., strong detrimental effects) by the majority of the operators, and the impact of fatigue was judged greatest on vigilance and overall operator effectiveness. They reported that fatigue was a result of both long work hours and poor sleep quality. The fatigue they were concerned with however, was that of "mental" fatigue, and not due to prolonged physical exertion. Others have reported on the effects of fatigue and their impact on task performance (Eysenck, 1983; Grandjean and Koji, 1971; Poulton, 1973).

These findings carry strong implications for airline passenger security screeners because of the similarity between tasks employed and the adverse consequences on operator performance. Given the perceived strong impact of boredom/monotony and fatigue, it would be desirable to develop a list of countermeasures that would help minimize these stress-related effects. This area appears to be fruitful for research opportunities where the potential findings could be useful for numerous industrial situations where monitoring plays a major role (e.g., nuclear power plants, air traffic control, agriculture inspection).

4.   HUMAN FACTORS ISSUES IN AIRLINE PASSENGER SECURITY SCREENING CHECKPOINT WORKSTATION DESIGN.

The goals of this research program are to define the abilities and traits required of successful airline passenger security screeners and to develop valid selection protocols for those abilities and traits. However, in the course of conducting the JTA it became readily apparent that several human factors issues regarding workstation design needed to be addressed. Inadequacies in workstation design can adversely impact screener performance and reduce the efficiency of the overall security procedures. Poor work environments may also contribute to fatigue and may affect personnel retention. Supporting this concept, Astley and Fox (1975) noted that providing an improved working environment—as a result of physical changes (e.g., ergonomically designed chairs, padded floor mats)—will influence personnel turnover and absenteeism in a positive direction. These issues are presented in brief since they are beyond the scope of the current work.

31

4.1 LIGHTING AND GLARE.

No standardization for illumination levels or placement of light sources was observed among the 15 security checkpoints visited. According to Vine (1982), dim lighting encourages slow movement and lethargic labor. Vine noted that simply replacing dark colors with bright shades and increasing lighting levels to at least 100 foot-candles is reported to have improved processing rates the movement of people by 30 percent.

At two checkpoints, both located in an atrium area with glass canopies, severe glare from sunlight was a problem at the X-ray screening positions. The image on the display monitors could not be adequately seen. The degree of glare was severe enough to adversely impact screener performance. At both of these checkpoints, screeners used pyramid-shaped shields attached to the face of the display monitors. Although successfully eliminating the glare, these devices had apertures that sharply restricted the operator's field of vision. Further design work is needed to eliminate glare problems without affecting screener performance.

4.2 FLOOR GRADE.

One checkpoint was noted to have an estimated grade of 3 percent over the length of the checkpoint area. All personnel, whether seated or prone, had to perform the job on an uneven surface. The effects on performance, or contribution to fatigue, need to be addressed in such a work environment. Optimally, this checkpoint could be reconfigured to eliminate the uneven floor grade.

4.3 WORKSTATION ARCHITECTURE.

Almost 20 years ago, Astley and Fox (1975) indicated that a major ergonomics commitment was needed to redesign the physical dimensions of the industrial inspection workplace using anthropometric and biomechanical data. Nonetheless, several workstations were constructed such that the screener's view of approaching passengers was obstructed. In some cases the obstruction was caused by placement of the screening units next to permanent structures (i.e., walls, support pillars) of the facility, but in others, the obstruction was caused by the actual design of the checkpoint station.

This was particularly true of the so-called "high technology" checkpoints. High technology checkpoints often had portable walls placed between the X-ray position and the entry side of the screening unit. The view of approaching individuals was completely obstructed in most cases. This disadvantage requires that either another screener orally initiate activating the conveyor belt, or that the belt is run continuously during the shift. Typically the screening unit was kept continuously operational. Continuous operation demands that the screener maintain vigilance on the display monitor throughout the shift. The literature is rich with data that demonstrates vigilance declines over a 30-minute interval. It is critical to ensure that constant vigilance is not required over 30-minute intervals.

Allowing the view of oncoming passengers permits a screener to activate the X-ray screening device only when needed to screen carry-ons. This permits the screener numerous periods of inactivity (i.e., rests) and decomposes the shift into a multitude of short sessions of operations. Vigilance is then only required during brief periods of operations. Additionally, the screener has direct expectations of when to maintain vigilance. Requiring screeners to maintain constant vigilance on the display over the entire shift could easily cause partial or full images to pass by unseen.

The recommended solutions to these design flaws are to: (a) replace opaque walls with transparent Plexiglas structures; (b) install convex mirrors near the entry portals to permit view of oncoming passengers; or 3) install weight-activated micro-switches on the conveyor belts to activate the X-ray scanning device, with a visual or auditory cue to alert the screener to device activation. The first two solutions are cost-effective and would reduce screener workload.

A similar problem was noted with regard to the design of bag check stations. At some high technology checkpoints portable partitions obstructed the view of bag check stations from the X-ray station. This design flaw prevents screeners from seeing the results of a requested bag check. An important element of any training program is the use of feedback to develop expertise and proficiency. This is vital in developing X-ray scanning proficiency in screener personnel. Feedback is necessary for screeners to develop skills in identifying objects from their X-ray images. Repetition in relating X-ray images to the visual sight of objects is an important tool for skill development. When this feedback loop is removed an important training technique is lost. It is recommended that bag check stations are always in full view of the X-ray scanning position to make available the opportunity for training.

4.4   COMMUNICATION AND DATA TRANSFER SYSTEMS.

High technology checkpoints characteristically linked X-ray screeners with CSS personnel using headsets. An additional feature of these checkpoints includes that transmission of the X-ray images from each checkpoint X-ray display monitor to a central information center staffed by the CSS. We observed that the CSS could be viewing as many as four X-ray display monitors simultaneously on four individual displays. These monitors could also display information from remote video cameras, and not necessarily be receiving images from the operational X-ray monitors. Screeners operating the X-ray scanning position were not in visual contact with the CSS and would not be aware of what, if any, information was being received by the CSS from their individual monitor.

The consequences of these design features are multi-faceted. Issues regarding the diffusion of responsibility for screening carry-ons need to be addressed to determine if X-ray screeners are less vigilant in their performance if they believe a supervisor is also screening from a remote location. The impact of isolating the CSS from screeners also needs to be explored with regards to the affects on training effectiveness, team cohesiveness, supervisory contact, motivation, and job competence. We observed that the visual isolation created by the architecture of these systems was not conducive to building a team approach to security. X-ray positions were isolated from viewing approaching passengers, bag check stations, and visual contact with

supervisors. In short, the X-ray screener was limited in their situational awareness of the entire security environment.

4.5   FURTHER CONSIDERATIONS.

Several additional features of airline security screening checkpoints were found to warrant human factors engineering intervention (see table 5). Among these were the need to address: (a) the seating provided for the X-ray scanning position; (b) the ambient noise levels around checkpoints and the need for sound absorption materials; (c) the effect of night work and circadian rhythms on screening performance; (d) the need for training of handling and lifting heavy loads; and (e) good management training to supervisors (e.g., TQM). These are only a few of the issues that warrant attention. The investigator is referred to Grandjean's (1980) "Checklist for the Analysis of the Workplace" and MIL-STD-1472C "Human Engineering Design Criteria for Military Systems, Equipment and Facilities" for a comprehensive treatment.

TABLE 5. WORKSTATION EVALUATION AT SAN FRANCISCO INTERNATIONAL AIRPORT (SFO)

	CONCOURSE					
	A	B	C	D	E	F
Floor grade	Y (steep)	Y (2 percent)	Y	Y (2 percent)	N	N
Noise level	N	N	N	N	N[a]	N
Glare problem(s)	N	N	N[b]	N	N[b]	N
Lighting level	Low	N	N	N	Low[c]	N
Visibility/Obstructions	N	N	N	N	N	N
Distance to exit	N	N	N	N	N[a]	Y (30 yds)
Space	Y	N	N	N	N	N
2-Stage Mag	N	N	N	Y	N	Y
"Read" position	Stand	Stand	Stand	Stand	Stand	Sit
Climate	N	N	N	N	N	N

(Y = problem, N = no problem)

[a] Moderate level of vibration under stations by joiner beam that runs width of checkpoint; high frequency nature.
[b] Late afternoon through rotunda.
[c] Glare by type of lighting installed.

These issues were presented to highlight some of the problematic areas related to security checkpoint design. The significance of these factors is a concern for the impact they present on human performance. In a work environment where fatigue and vigilance decrements can rapidly degrade performance and system efficiency, these issues present challenges that require attention.

It is also plausible that a poor work environment may contribute to decreased job longevity among screeners. Most of the issues presented can be resolved with minimal investment. We agree with other investigators (Drury and Fox, 1975) who stated, "... inspection has proved an exciting vehicle for ergonomics" (p. 299). Their observation is quite accurate based upon our initial findings.

5.   OTHER SOURCES.

1.   Adams, J.A., Humes, J.M. and Stenson, H.H. (1962). Monitoring of Complex Visual Displays: III. Effects of Repeated Sessions on Human Vigilance. Human Factors, 4: 149–157.

2.   Astley, R.W., and Fox, J.G. (1975). Human Reliability in Quality Control. The Analysis of an Inspection in the Rubber Industry. C.G. Drury and J.G. Fox (Eds.). New York: Taylor and Francis, pp. 253–272.

3.   Atkinson, A.P.C. (1973). Selection of the Necessary but Not Sufficient Skills for a Job. Human Factors, 15(2): 125–128.

4.   Bakan, P., Belton, J.A., and Toth, J.C. (1963). Vigilance: A Symposium. Extraversion-Introversion and Decrement in an Auditory Vigilance Task. D.N. Buckner and J.J. McGrath (Eds.). New York: McGraw-Hill.

5.   Berch, D.B., and Kanter, D.R. (1984). Sustained Attention in Human Performance J. Warm (Ed.). Chichester, UK: Wiley, pp. 144–178.

6.   Bevan, W., and Turner, E.O. (1965). Vigilance Performance with a Qualitative Shift in Verbal Reinforcers. Journal of Experimental Psychology, 70: 83–86.

7.   Boff, K.R., and Lincoln, J.E. (1988). Engineering Data Compendium: Human Perception and Performance. AAMRL, Wright-Patterson AFB, OH.

8.   Brennan, P.J. (1974). Job Satisfaction: Is There a Trend? Washington, DC: U.S. Government Printing Office.

9.   Broadbent, D.E. (1958). Perception and Communication. New York: Pergamon.

10.   Buckner, D.N., and McGrath, J.J. (1963). Vigilance: A Symposium. A Comparison of Performance on Single and Dual Sensory Mode Vigilance Tasks. D.N. Buckner and J.J. McGrath (Eds.). New York: McGraw-Hill.

11.   Carmody, D.P. (1985). Eye Movements and Human Information Processing. Free Search, Restricted Search, and the Need for Context in Radiologic Image Perception. R. Groner, G.W. McConkie, and C. Menz (Eds.). Elsevier, North Holland, pp. 341–355.

12.   Childs, J.M. (1976). Signal Complexity, Response Complexity, and Signal Specification in Vigilance. Human Factors, 18(2): 149–160.

13.   Coninsx, M. (1994). Seminar in Asia/Pacific Region Focuses on Role of Human Factors in Aviation Security. Knowles, A. (Ed.). ICAO Journal, 49: 16–18.

14. Craig, A. (1984). Sustained Attention in Human Performance. Human Engineering: The Control of Vigilance. J. Warm (Ed.). Chichester, UK: Wiley, pp. 247–291.

15. Craig, A., and Colquhoun, W.P. (1975). Human Reliability in Quality Control. Vigilance: A Review. C.G. Drury and J.G. Fox (Eds.). New York: Taylor and Francis.

16. Craig, A., and Colquhoun, W.P. (1977). Vigilance: Theory, Operational Performance, and Physiological Correlates. Vigilance Effects in Complex Inspection. R.R. Mackie (Ed.). New York: Plenum, pp. 239–262.

17. Crumley, L.M., Pierce, L.G., Schwalm, R.C., Coke, J.S., and Brown, J.C. (1992). Predicting Target Detection Performance Using the Armed Services Vocational Aptitude Battery Subtests and Cognitive Factor Tests. (ARI Research Note 92-56). Alexandria, VA: U.S. Army Research Institute.

18. Davies, D.R. (1979). Search and the Human Observer. Theories of Vigilance and Sustained Attention. M.A., J.N. Clare (Eds.). London: Taylor and Francis, pp. 14–19.

19. Davies, D.R., and Parasuraman, R. (1982). The Psychology of Vigilance. Academic Press: London.

20. Davies, D.R., and Tune, G.S. (1970). Human Vigilance Performance. London: Staples Press.

21. Drury, C.G., and Fox, J.G. (1975). Human Reliability in Quality Control. The Inspector in the Future: His Efficiency and Well-Being. C.G. Drury and J.G. Fox (Eds.). New York: Taylor and Francis, pp. 299–301.

22. Drury, C.G., and Sheehan, J.J. (1969). Ergonomic and Economic Facts—an Industrial Inspection Task. International Journal of Production Research, 7: 333–341.

23. Edwards, D. (1990). Pilot Mental and Physical Performance. Ames, IA: Iowa State.

24. Eysenck, M.W. (1982). Attention and Arousal: Cognition and Performance. Berlin: Springer-Verlag.

25. Eysenck, M.W. (1983). Stress and Fatigue in Human Performance. Incentives. G.R.J. Hockey (Ed.). Chichester: Wiley.

26. Federal Aviation Administration. (1991, February). Semiannual Report to Congress on the Effectiveness of the Civil Aviation Security Program. (Tech. Rep. DOT/FAA/ACS-89-1[31]). Washington, DC: NTIS.

27. Fox, J.G. (1975). Human Reliability in Quality Control. Vigilance and Arousal: A Key to Maintaining Inspectors' Performance. C.G. Drury and J.G. Fox (Eds.). New York: Taylor and Francis.

28. Frasher, R.G., and Pare, J.A.P. (1977). Diagnosis of Diseases of the Chest. Philadelphia: Saunders.

29. Funke, D.J. (1979, September). Security Management. Perceptual Training Sharpens Baggage Inspection Skills, pp. 30–34.

30. General Accounting Office. (1987a, April). FAA Preboard Passenger Screening Test Results. (GAO/RCED-87-125FS): Washington, DC.

31. General Accounting Office. (1987b, July). Aviation Security: FAA Needs Preboard Passenger Screening Performance Standards. (GAO/RCED-87-182): Washington, DC.

32. Grandjean, E., and Koji, K. (1971). Methodology in Human Fatigue Assessment. Introductory Remarks. K. Hashimoto, K. Koji, and E. Grandjean (Eds.). London: Taylor and Francis.

33. Hertzberg, F. (1966). Work and the Nature of Man. Cleveland: World Publishing.

34. Howell, W.C., Johnston, W.A., and Goldstein, I.L. (1966). Organizational Behavior and Human Performance. Complex Monitoring and Its Relation to the Classical Problem of Vigilance, 1: 129–150.

35. Howland, D., and Wiener, E.L. (1963). Vigilance: A Symposium. The System Monitor. D.N. Buckner and J.J. McGrath (Eds.). New York: McGraw-Hill, pp. 217–223.

36. Hughes, D. (1993, June 7). Aviation Week and Space Technology. FAA Prepares Testing Procedure for Explosive Detection Systems, pp. 122–125.

37. ITS (1993, April). The Personnel Performance Study. ITS Training and Personnel Development Department. Cleveland, OH: Author.

38. Jerison, H.J., and Wing, J.F. (1963). Vigilance: A Symposium. Human Vigilance and Operant Behavior. D.N. Buckner and J.J. McGrath (Eds.). New York: McGraw-Hill, pp. 34–38.

39. Kaempf, G., Klinger, D., and Wolf, S. (1994). Development of Decision-Centered Interventions for Airport Security Checkpoints. (Technical Report: DTRS-57-93-C-00129). Fairborn, OH: Klein Associates, Inc.

40. Keeping the Operator in Line. (1992, June). Airport Support, p. 41.

41. Kelly, M.L. (1955). A Study of Industrial Inspection by the Method of Paired Comparisons. Psychological Monographs, 69: 1–16.

42. Kjellberg, A. (1977). Vigilance: Theory, Operational Performance, and Physiological Correlates. Sleep Deprivation, Arousal, and Performance. R.R. Mackie (Ed.). New York: Plenum.

43. Krupski, A., Raskin, D.C., and Bakan, P. (1971). Physiological and Personality Correlates of Commission Error in an Auditory Vigilance Task. Psychophysiology, 8: 304–311.

44. Levine, J.M. (1966). The Effects of Values and Costs on the Detection and Identification of Signals in Auditory Vigilance. Human Factors, 8: 525–537.

45. Linestone, H. (1975). The Delphi Method: Techniques and Applications. Eight Basic Pitfalls: A Checklist. H. Linstone, and M. Turoff (Ed.). New York: Addison Wessley Publishing.

46. Loeb, M., and Alluisi, E.A. (1984). Sustained Attention in Human Performance. Theories of Vigilance. J. Warm (Ed.). Chichester, UK: Wiley, pp. 179–205.

47. Mackie, R.R. (1987). Vigilance Research—Are We Ready for Countermeasures? Human Factors, 29(6): 707–723.

48. Mackworth, N.H. (1950). Researches on the Measurement on Human Performance. (MRC Report Series No. 268, Cambridge).

49. Mackworth, N.H. (1957). Some Factors Affecting Vigilance. Advanced Sciences, 53: 389–393.

50. Mackworth, J.F. (1965). Effect of Amphetamine on the Detectability of Signals in a Vigilance Task. Canadian Journal of Psychology, 19(2): 104–110.

51. McGrath, J.J., Harabedian, A., and Buckner, D.N. (1960). Human Factors Problems in Anti-Submarine Warfare. (Tech. Rep. No. 4). An Exploratory Study of the Correlates of Vigilance Performance. Los Angeles: Human Factors Research.

52. Meister, D. (1989). Conceptual Aspects of Human Factors. Baltimore, MD: John Hopkins University Press.

53. Moder, J.J., and Oswalt, J.H. (1959). An Investigation of Some Factors Affecting the Hand Picking of Small Objects. Journal of Industrial Engineering, 10: 213–218.

54. Murrell, K.F. (1965). Human Performance in Industry. New York,: Reinhold, Inc.

55.   Nachreiner, F. (1977). <u>Vigilance: Theory, Operational Performance, and Physiological Correlates</u>. Experiments on the Validity of Vigilance Experiments. R.R. Mackie (Ed.). New York: Plenum, pp. 665–678.

56.   Nadler, E.D., and Mengert, P. (1993). <u>Safe Passage Data Analysis: Interim Report</u>. (Tech. Rep. No. DOT-VNTSC-FA3T5-PM-93-3). Cambridge, MA: DOT-FAA.

57.   Nelson, J.B., and Barany, J.W. (1969). A Dynamic Visual Recognition Test for Paced Inspection Tasks. <u>Transactions of Industrial Engineering</u>, 1: 327–332.

58.   Oster, C.V., Strong, J.S., and Zorn, C.K. (1992). <u>Why Airplanes Crash</u>. New York: Oxford.

59.   Parasuraman, R. (1986). <u>Handbook of Perception and Human Performance: Vol. II: Cognitive Processes and Performance</u>. Vigilance, Monitoring, and Search. K.R. Boff, L. Kaufman, and J.P. Thomas (Eds.). New York: Wiley, pp. 1–39.

60.   Pierce, L.G., Crumley, L.M., and Clifford, P.A. (1991). Detecting People Who Can Detect Targets. <u>Proceedings of the Human Factors Society 35th Annual Meeting</u>, 1005–1009.

61.   Potter, A.N. (1975). Is Anti-Hijacking Success Only Temporary? <u>Security World</u>, 12: p. 34.

62.   Poulton, E.C. (1973). The Effects of Fatigue Upon Inspection Work. <u>Applied Ergonomics</u>, 4: 73–83.

63.   President's Commission. (1990). <u>Report to the President by the President's Commission on Aviation Security and Terrorism</u>. Executive Order 12686, August 4, 1989: Washington, DC. Sheehan, J.J., and Drury, C.G. (1971). The Analysis of Industrial Inspection. <u>Applied Ergonomics</u>, 2: 74–78.

64.   Sipowicz, R.R., Ware, J., and Baker, R.A. (1962). The Effects of Reward and Knowledge of Results on the Performance of a Simple Vigilance Task. <u>Journal of Experimental Psychology</u>, 64: 58–61.

65.   Smith, R.L., Lucaccini, L.F., and Epstein, M.H. (1967). The Effects of Monetary Rewards and Punishments on Vigilance Performance. <u>Journal of Applied Psychology</u>, 51: 411–416.

66.   Smith, R.L., and Lucaccini, L.F. (1969). Vigilance Research: It's Application to Industrial Problems. <u>Human Factors</u>, 11: 149–156.

67.   St. John, P. (1991). <u>Air Piracy, Airport Security, and International Terrorism</u>. New York: Quorum Books.

68.   Study Tallies Hostile Acts against Civil Aviation. (1991, October). <u>Air Line Pilot</u>, p. 48.

69.   Sutton, D. (1969). <u>A Textbook of Radiology</u>. Baltimore, MD: Williams and Wilkins.

70.   Thackray, R.I., Bailey, J.P., and Touchstone, R. (1977). <u>The Effect of Increased Monitoring Load on Vigilance Performance Using a Simulated Radar Display</u>. (Rep. No. FAA-AM-77-18). Federal Aviation Administration, Washington, DC: Office of Aviation Medicine.

71.   Thackray, R.I., Jones, K.N., and Touchstone, R.M. (1974). Personality and Physiological Correlates of Performance Decrement on a Monotonous Task Requiring Sustained Attention. <u>British Journal of Psychology</u>, 65: 351–358.

72.   Thaher, F.F. (1991). <u>A Study of Security Safeguards at King Abdulaziz International Airport in Jeddah, Saudi Arabia</u>. Unpublished Graduate Research Project, Embry-Riddle Aeronautical University, Daytona Beach, FL.

73.   Tune, G.S. (1966). Errors of Commission as a Function of Age and Temperament in a Type of Vigilance Task. <u>Journal of Experimental Psychology</u>, 18: 358–361.

74.   U.S. Department of Defense. (1981). <u>Human Engineering Design Criteria for Military Systems, Equipment and Facilities</u>. (MIL-STD-1472C). Redstone Arsenal, AL: Author.

75.   Vickers, D., Nettelbeck, T., and Willson, R.J. (1972). Perceptual Indices of Performance: The Measurement of "Inspection Time" and "Noise" in the Visual System. <u>Perception</u>, 1: 262–295.

76.   Vine, D. (1982, November). <u>Airport Services Management</u>. Airport Security Tips (For Those New to the Screening Game). pp. 24–25.

77.   Vosburgh, A.R. (1993). <u>Analysis of Some Aspects of Aviation Security Today Considered against Factors in the 1988 Bombing of Pan Am Flight 103</u>. Unpublished Graduate Research Project, Embry-Riddle Aeronautical University, Daytona Beach, FL.

78.   Wallis, R. (1992, June). <u>Jane's Airport Review</u>. U.S. Proposes Criminal Background Checks, pp. 21–26.

79.   Warm, J.S., Dember, W.N., Lanzetta, T.M., Bowers, J.C., and Lysaght, R.J. (1985). Information Processing in Vigilance Performance: Complexity Revisited. <u>Proceedings of the Human Factors Society 29th Annual Meeting</u>, 1: 19–22.

80.   Wickens, C.D. (1992). <u>Engineering Psychology and Human Performance</u>, 2nd edition. New York: Harper Collins.

81. Wiener, E.L. (1969). Money and the Monitor. Perceptual and Motor Skills, 29: 627–634.

82. Wiener, E.L. (1975). Human Reliability in Quality Control. Individual and Group Differences in Inspection. C.G. Drury and J.G. Fox (Eds.). London: Taylor and Francis, pp. 101–122.

83. Wiener, E.L. (1987). Application of Vigilance Research: Rare, Medium, or Well Done? Human Factors, 29(6): 725–736.

84. Williams, L.G. (1966). Target Conspicuity and Visual Search. Human Factors, 8(2): 80–91.

85. Wylie, C.D., Mackie, R.R., and Smith, M.J. (1985). Stress and Sonar Operator Performance: A Literature Review and Operator Survey. (Tech. Rep. 483-1). Goleta, CA: Essex Corp.